

TARTU ÜLIKOOL
MATEMAATIKA-INFORMAATIKATEADUSKOND
MATEMAATILISE STATISTIKA INSTITUUT

Kristi Tüli

Võrdeliste tõenäosustega valik tervishoiutöötajate tunnipalga uurimiseks

Bakalaureusetöö (6 EAP)

Juhendaja:
Natalja Lepik, PhD

TARTU
2014

Võrdeliste tõenäosustega valik tervishoiutöötajate tunnipalga uurimiseks

Käesoleva bakalaureusetöö eesmärk on võrrelda nelja valimi moodustamise viisi, mis on välja pakutud Tervise Arengu Instituudi poolt ning vastab teatud kriteeriumitele. Töö praktilises pooles viikase läbi simulatsioon 2013. aasta andmetega, et uurida huvipakkuva tunnuse parameeterhinnangute täpsust. Teoreetilises osas tutvustatakse tõenäosuslike valikuuringutega seotuid mõisteid, valimi võtmise viise ja hinnanguteks vajalikke valemeid. Hinnangud leitakse osakogumites. Käesolevas töös kasutatakse lihtsat juhuslikku valikut, suurusega võrdelist tõenäosuslikku valikut ning kihtvalikut. Leitakse ka kitsendustega hinnangud kogusummadele ja seejärel võrreldakse saadud hinnanguid ja antakse ülevaade, mis oleks parim valiku võtmise meetod.

Märksõnad: kihtvalik, lihtne juhuslik valik, osakogumid, suurustega võrdeline tõenäosuslik valik, tõenäosuslikud valikud, valikuteooria, valikuuringud.

Probability proportional to size sampling to study healthcare personnel's hourly wage

The purpose of this thesis is to compare four ways of constructing a sample proposed by The National Institute for Health Development that meets certain criteria. A simulation is conducted in the practical part of this thesis using data from 2013 to analyze parameter estimation accuracy of variable of interest. In theoretical part, definitions of probability sampling surveys, ways and formulas of sampling are introduced. Estimations are calculated in subsamples. Simple random sampling, probability proportional to size sampling and stratified sampling are used in this thesis. General restriction estimations for total sum are calculated and after compared with each other, an overview of the best sampling method is given.

Keywords: stratified sampling, Simple Random Sampling, population subsamples, probability proportional to size sampling, probability sampling, sampling theory, sample survey.

Sisukord

1.	Sissejuhatus	4
2.	Tõenäosuslik valikuuring	6
3.	Valikudisain ja selle karakteristikud	7
3.1	Valikudisaini karakteristikud.....	7
3.2	Lihtne juhuslik valik tagasipanekuta	8
3.3	Suurusega võrdelise tõenäosusega valik tagasipanekuga.....	9
3.4	Kihtvalik	9
4.	Hindamine	11
4.1	Keskmise π - ja p -hinnangud	12
5.	Osakogumid ja keskmise hindamine osakogumites	14
5.1	LJV TTA korral keskmise hindamine osakogumites	15
5.2	PPS tagasipanekuga valiku korral keskmise hindamine osakogumites.....	16
6.	Kitsendustega hindamine	18
6.1	GR - hinnang	18
7.	Hinnangute täpsusnäitajad.....	21
8.	Simuleerimisülesanne.....	23
8.1	Andmestiku kirjeldus.....	23
8.2	Valimi võtmine ja parameetrite hindamine	27
8.2.1	Keskmise põhitunnipalga hinnang LJV TTA korral	27
8.2.2	Keskmise põhitunnipalga hinnang PPS TGA korral.....	29
8.2.3	Keskmine põhitunnipalga hinnang lihtsa juhusliku kihtvaliku korral	30
8.2.4	Keskmine põhitunnipalga hinnang PPS TGA kihtvaliku korral	32
8.2.5	Kitsendustega hinnang	35
9.	Kokkuvõte	37
10.	Kasutatud kirjandus.....	39
11.	Lisad	40
	Lisa 1.....	40
	Lisa 2	41
	Lisa 3	42
	Lisa 4.....	48
	Lisa 5	54
	Lisa 6	59
	Lisa 7	63

1. Sissejuhatus

Tervis on tänapäeval väga kõrgelt hinnatud, samuti tervishoiutöötajad. Tervishoiuasutuste töötajatel on kõrge vastutusega töökoht, sellest tulenevalt võiks neil olla ka vastav palk. Paraku räägitakse tihti, kuidas meditsiinitöötajad lahkuvad Eestist, kuna neile ei maksta väärilist palka. Tervise Arengu Instituut (TAI) uurib iga aasta tervishoiutöötajate tunnipalka ning sellega seonduvaid andmeid. Üldkogumi parameetrite välja selgitamiseks on erinevaid võimalusi. Näiteks teostada kõikne uuring ja seejärel teha järeldused huvipakkuvate tunnuste kohta või töötada välja valikuuring, mis annaks soovitud täpsusega hinnangud meid huvitavatele parameetritele. Kõikse uuringu miinuseks on kõrge hind ja töömahuks. Hetkel teostab TAI kõikset uuringut, et väljastada tervishoiutöötajate tunnipalgaga seotud andmeid. See on kallis ja töömahukas. TAI soovib leida sobivat valiku võtmise meetodit, mis annaks huvipakkuvatele parameetritele hinnangud heade täpsustega.

Käesoleva töö eesmärk on võrrelda kahte valikuuringut, mida võiks TAI oma valimi välja töötamiseks kasutada. Nende kahe valiku võtmise meetodi tulemusi võrreldakse lihtsa juhusliku tagasipanekuta disainiga ja lihtsa juhusliku kihtvalikuga. Samuti rakendatakse valikudisainilt saadud hinnangutele uudset meetodit – kitsendustega hinnangut. See ühelt poolt vähendab üldjuhul hinnangute varieeruvust, ja teiselt poolt, võimaldab saavutada hinnangute jaoks teatud kitsenduste täitmist. Parameeterhinnangud leitakse osakogumites. Üldkogum jaotakse osakogumiteks kahte erinevat moodi. Esiteks tervishoiuteenuse osutaja liigi järgi ning teiseks ametigruppide järgi.

Töö esimeses pooles antakse põgus ülevaade tõenäosuslikest valikuuringutest, disainidest, parameetrite hindamisest ja nendega seonduvatest mõistetest. Tuuakse välja valemid, mida huvipakkuva parameetri hindamiseks kasutatakse ning räägitakse töös kasutatavatest hinnangute täpsusnäitajatest. Samuti selgitatakse osakogumi mõistet ning neis hinnangute leidmist.

Töö teises pooles teostatakse Tervise Arengu Instituudilt saadud andmete peal valimi võtmise meetodeid. Simuleeritakse 1000 valimit ning nende pealt hinnatakse põhitunnipalka. Väljastatakse ja interpreteeritakse nelja erineva valikudisaini korral saadud tulemused ning nende täpsusnäitajad. Samuti esitatakse tulemused, mis on saadud pärast kitsenduste rakendamist. Analüüsitakse, mis oleks parim ehk täpsem meetod valiku tegemiseks.

Bakalaureusetöö on kirja pandud ja vormistatud tekstitöötlustarkvaras Microsoft Word 2013. Andmete analüüs on tehtud statistikatarkvaras SAS. Kitsendustega hinnangud on leitud statistikapaketi R abil. Programmi kood on esitatud lisades 3-7. Tulemused koguti ja vormistati tabelitöötlusprogrammiga Microsoft Excel 2013.

Autor tänab käesoleva bakalaureusetöö juhendajata Natalja Lepikut arvukate nõuannete ja näpunäidete eest. Samuti tänab Tervise Arengu Instituuti andmete kasutamise loa eest.

2. Tõenäosuslik valikuuring

Vaikuuring on statistiline uuring, milles otsustused üldkogumi kohta tehakse valimi (üldkogumi ühe osa) baasil (Traat, Inno, 1997, lk 13). Ainult valimisse kuuluvatelt objektidelt kogutakse infot ning nende põhjal hinnatakse üldkogumi parameetreid. Valikuuringute teostamine on tänapäeval üsna levinud, väga palju kasutatakse seda majandus- ja sotsiaalvaldkondades uuringute tegemisel. Valikuuringute eelisteks peetakse väiksemat maksumust, suuremat kiirust ja täpsust. Väiksem maksumus ning suurem kiirus tulenevad uuritavate objektide väiksemast arvust ning oskuslikust valimi planeerimisest.

Valimi võtmise meetodeid saab oma ülesehituse ja kasutusvaldkonna järgi jagada kahte põhilisse klassi: empiirilised ja tõenäosuslikud valikud. Käesolevas bakalaureusetöös kasutatakse tõenäosuslike valikumeetodeid. Tõenäosusliku valiku korral on iga üldkogumi objekti kohta teada tema valimisse sattumise tõenäosus. Objekti valik üldkogumist on juhuslik, aga sõltub siiski tema tõenäosusest sattuda valimisse. Traat, Inno (1997) raamatus lk 57 on tõenäosuslik valik defineeritud selliselt, et valik üldkogumist, mille korral

- on võimalik defineerida kõigi võimalike valimite hulga

$$S = \{s_1, s_2, \dots, s_M\};$$

- iga valimi $s \in S$ jaoks on teada tema valikutõenäosus $p(s)$;
- iga üldkogumi objekti valimisse sattumise tõenäosus on teada ja on positiivne;
- valmi võtmiseks kasutatav juhuslik mehhanism tagab, et valimi s valikutõenäosus on $p(s)$.

Objektide juhuslikkust valimisse sattumisel saab koordineerida mitmel viisil, näiteks, tagasipanekuga (TGA) ja tagasipanekuta (TTA) valik. Käesolevas töös rakendatakse TGA ja TTA valikuid. Tagasipanekuta valiku korral ei panda valimisse sattunud objekti üldkogumisse tagasi, seega järgnevatel valikusammudel seda objekti enam valida ei saa. Tagasipanekuga valiku korral kantakse valimisse sattunud objekti andmed valimisse, aga üldkogumist ei eemaldata. Üks ja sama objekt võib osutuda valituks mitu korda, seega ka tema andmed võivad minna valimisse korduvalt. (Traat, Inno, 1997, lk 23)

3. Valikudisain ja selle karakteristikud

Andmete valimisse võtmiseks on palju erinevaid valikudisaine. Käesolevas bakalaureusetöös kasutatakse kolme erinevat valikudisaini – lihtne juhuslik valik tagasipanekuta (*Simple Random Sampling without Replacement*, eestikeelne lühend on LJV TTA, mida kasutatakse ka käesolevas töös), suurusega võrdelise tõenäosusega valik tagasipanekuga ja kihtvalik. Suurusega võrdelist tõenäosusega disaini nimetatakse ka PPS disainiks, mis tuleb inglise keelsest väljendist *probability proportional to size sampling*. Valikudisainil on oluline roll valikuteoorias, selle abil määratakse hinnangute statistilised omadused. Hinnangufunktsiooni moodustamiseks ning selle statistiliste omaduste uurimiseks ei kasutata otseselt valikudisaini ennast vaid selle poolt määratud karakteristikuid: objektide kaasamis- ja valikutõenäosusi.

Edasine valikudisainide ja disaini karakteristikute kirjeldus põhineb raamatul Traat, Inno (1997).

3.1 Valikudisaini karakteristikud

Järgnevas osas on antud ülevaade tähtsamatest mõistetest, mida kasutatakse disainipõhises valikuuringute teoorias.

Definitsioon 1. Valikudisainiks nimetatakse tõenäosusjaotust $p(s)$ kõigi antud valiku jaoks võimalike valimite hulgal S .

Valikudisainiga $p(s)$ on määratud kõigi hinnangute statistilised omadused.

Definitsioon 2. Üldkogumi objekti i ($i = 1, 2, \dots, N$) kaasamistõenäosuseks π_i nimetatakse tõenäosust, millega see objekt kaasatakse valimisse antud disaini $p(s)$ korral.

Definitsioon 3. Üldkogumi objekti i ($i = 1, 2, \dots, N$) valikutõenäosuseks p_i nimetatakse tõenäosust, millega seda objekti võidakse valida antud disaini ühel sammul.

Tõenäosused π_i ja p_i on teadaolevad suurused üldkogumi iga objekti jaoks ning nad on määratud valikudisainiga. Valikutõenäosust vaadatakse tavaliselt seoses tagasipanekuga disainidega.

Definitsioon 4. Valikuindikaator I_i on määratud iga üldkogumi objekti i ($i = 1, 2, \dots, N$) jaoks ja näitab i -nda objekti valikute arvu.

Tagasipanekuta disainide korral on see binaarne juhuslik suurus $I_i \sim \text{Be}(\pi_i)$

$$I_i = \begin{cases} 1, & \text{kui } i - \text{ndas objekt on valimis;} \\ 0, & \text{muidu.} \end{cases}$$

Valikuindikaatori keskvärtus, dispersioon ja kovaratsioon on tagasipanekuta valiku korral määratud kaasamistõenäosustega:

- $E(I_i) = \pi_i$
- $V(I_i) = \pi_i (1 - \pi_i)$
- $\text{Cov}(I_i, I_j) = \pi_{ij} - \pi_i \pi_j$,

kus $\pi_i = P(I_i = 1)$ ja $\pi_{ij} = P(I_i = 1, I_j = 1)$ on vastavalt esimest ja teist järku kaasamistõenäosused.

Tagasipanekuga valiku korral $I_i \sim \text{Bin}(n, p_i)$ ja karakteristikud on järgmised:

- $E(I_i) = np_i$
- $V(I_i) = np_i (1 - p_i)$
- $\text{Cov}(I_i, I_j) = - np_i p_j$.

3.2 Lihtne juhuslik valik tagasipanekuta

LJV TTA korral on kõigil üldkogumi objektidel võrdne tõenäosus osutada valituks. Lihtsa juhusliku valikuga teostatud valim peegeldab üldkogumi struktuuri ja on isekaaluv¹.

Olgu $U = \{1, \dots, N\}$ lõplik üldkogum. Üldkogumist U saab moodustada n mahulisi hulki kokku $C_N^n = \frac{N!}{(N-n)!n!}$. Nende hulkade hulk on lihtsa juhusliku valiku kõigi võimalike valimite hulk $S = \{s_1, \dots, s_M\}$, kus $M = C_N^n$. (Traat, Inno, 1997, lk 90-91)

Definitsioon 5. Lihtsa juhusliku valiku disainiks nimetatakse diskreetset pidevat jaotust $p(s)$ kõigi valimite hulgal S , kus $p(s) = 1/C_N^n$, kus $s \in S$.

LJV TTA karakteristikud:

- $E(I_i) = \frac{n}{N}$ ($=f$, mida nimetatakse valikusuhteks)
- $V(I_i) = \frac{n}{N} \left(1 - \frac{n}{N}\right)$
- $\text{Cov}(I_i, I_j) = \frac{n(n-1)}{N(N-1)} - \frac{n^2}{N^2} = -f \frac{1-f}{N-1}$

LJV TTA teostamise lihtne näide. Oletame, et üldkogum on ühe kooli neljanda klassi õpilased, keda on kokku N , ja tahetakse moodustada valim suurusega n . Kirjutatakse iga

¹ Isekaaluv valim – ühesuguse tähtsusega objektidest koosnev valim, iga objekt valimise esindab võrdset arvu üldkogumi objekte. Hinnangute arvutamisel isekaaluvalt valimilt ei ole objektidele vaja omistada erinevaid kaalusid. (Traat, Inno, 1997, lk 16)

õpilase nimi ühele paberlipikule ja pannakse kaabu sisse, seejärel võetakse pimesi tagasipanekuta kaabust n nimelipikut ja need, kelle nimed on paberil, satuvadki valimisse.

Lihtsat juhuslikku valikut kasutatakse praktikas üsna sageli, eelkõige sellepärast, et ta on teoreetiliselt väga hästi uuritud. Hinnangud ja hinnangute omadused on LJV korral sarnased klassikalise statistika valemitega. Tihti osutub LJV teoreetiliseks baasiks, millega võrreldakse teisi disaine. Samas LJV ei võimalda reguleerida objektide kaasamistõenäosusi, mistõttu ei saa ka reguleerida maksumust, tegemise kiirust ja uuritava tunnuse varieeruvusest tingitud viga.

3.3 Suurusega võrdelise tõenäosusega valik tagasipanekuga

Suurusega võrdelise tõenäosusega valiku korral objektide valik üldkogumist allub erinevatele valikutõenäosustele, mis on võrdelised objektide suurustega (või mingi teise objekti iseloomustava mõõduga). Selle tõttu satub valimisse suuremaid objekte rohkem. Tulemust rohkem mõjutavatele objektidele antakse suuremad valikutõenäosused ja vähem mõjutavatel väiksemad. Sellega on võimalik vähendada hinnangu varieeruvust ning suurendada täpsust.

Tagasipanekuga valiku korral määratakse valikutõenäosused võrdeliselt tausttunnusega x , st $p_i \propto x_i$. Tausttunnuse korral on oluline see, et peab teadma selle tunnuse väärtuseid kõikide objektide kohta üldkogumis ning see tunnus peab olema tugevalt korreleeritud uuritava tunnusega. Kuna $1 = \sum_U p_i$ ja $\sum_U x_i$ on seotud sama võrdeteguriga, siis

$$p_i = \frac{x_i}{\sum_U x_j}.$$

(Traat, Inno, 1997, lk 24, 123)

3.4 Kihtvalik

Kihtvaliku korral jagatakse üldkogum mingi teadaoleva tausttunnuse abil osadeks ehk kihtideks. Igat kihti võib käsitleda eraldi iseseisva kogumina. Kihtvalik on praktikas väga laialdaselt levinud disain, üheks põhjuseks on just see, et ta võimaldab kihtides rakendada erinevaid valikumeetodeid. Iga kihi valimis leitakse selle kihi parameetrite hinnangud ning neid sobivalt kombineerides leitakse üldkogumi hinnangud. Näiteks, kui üldkogumiks on kõik Eesti koolid, siis saame need maakondade järgi kihtidesse jagada. Tekib 15 kihti, kus igas neist saame võtta soovitud valikumeetodiga valimi ja seejärel leida soovitud hinnangud.

Olgu lõplik üldkogum $U = \{1, \dots, i, \dots, N\}$ jagatud H mittelõikuvaks kihiks $U_1, \dots, U_h, \dots, U_H$, kus U_h tähistab kihi h kõigi objektide hulka. Igas kihis U_h toimub valik sõltumatult teistest

kihtidest ja rakendatakse disaini $p_h(s)$. Kogu valim moodustub kihtide valimite ühendina. Kuna valikud on sõltumatud, siis valimi s saamise tõenäosus avaldub valikudisainide tõenäosuste korrutisena $p(s) = p_1(s)p_2(s)\dots p_H(s)$. Üldkogumi ja valimi mahud avalduvad järgmiste summadena:

$$N = \sum_{h=1}^H N_h$$

$$n = \sum_{h=1}^H n_h ,$$

kus N_h on objektide arv kihis h ja n_h on valimi objektide arv kihis h .

4. Hindamine

Selles peatükis tutvustatakse edaspidiseks vajaliku üldist hindamisteoreemi ja esitatakse sellest lähtuvalt hinnang üldkogumi keskmisele.

Üldine hindamisteoreem. Üldkogumi kogusumma $t = \sum_U y_i$ nihketa² hinnang on

$$\hat{t} = \sum_U I_i \check{y}_i \text{ (või } \hat{t} = \sum_U \omega_i y_i \text{),}$$

kus $\check{y}_i = \frac{y_i}{E(I_i)}$ ja $\omega_i = \frac{I_i}{E(I_i)}$ on disaini kaal.

Selle hinnangu disainipõhine dispersioon on

$$V(\hat{t}) = \sum \sum_U \Delta_{ij} \check{y}_i \check{y}_j,$$

kus $\Delta_{ij} = \text{Cov}(I_i, I_j)$. Dispersiooni nihketa hinnanguks $E(I_i I_j) > 0$ korral on

$$\hat{V}(\hat{t}) = \sum \sum_U \check{\Delta}_{ij} \check{y}_i \check{y}_j I_i I_j \text{ (või } \hat{V}(\hat{t}) = \sum \sum_U \check{\Delta}_{ij} \omega_i y_i \omega_j y_j \text{),}$$

kus $\check{\Delta}_{ij} = \frac{\Delta_{ij}}{E(I_i I_j)}$.

Üldine hindamisteoreem kehtib nii TTA kui TGA valikudisainide korral. Vaja on teada vaid disainikarakteristikuid – $E(I_i)$, $E(I_i I_j)$ ja Δ_{ij} kui $i = j$ ja $i \neq j$ (Lepik, Traat, 2013, lk 27.)

Teoreemi tõestus kuulub aine „Valikuuringute teooria I“ programmi ja seetõttu siin ei esitata.

Üldkogumi keskmise \bar{Y} hinnangut saab moodustada kogusumma $t = \sum_U y_i$ hinnangu abil, kuna $\bar{Y} = \frac{t}{N}$.

Kui N on teada, saame üldisest hindamisteoreemist lähtudes nihketa hinnangu üldkogumi keskmisele:

$$\hat{\bar{Y}}_{\text{OHT}} = \frac{\hat{t}}{N}. \quad (1)$$

Dispersioon ja dispersiooni hinnang lähtuvad samuti üldisest hindamisteoreemist:

$$V(\hat{\bar{Y}}) = V\left(\frac{\hat{t}}{N}\right) = \frac{1}{N^2} V(\hat{t})$$

$$\hat{V}(\hat{\bar{Y}}) = \hat{V}\left(\frac{\hat{t}}{N}\right) = \frac{1}{N^2} \hat{V}(\hat{t})$$

(Lepik, Traat, 2013, lk 32).

² Hinnang $\hat{\theta}$ on parameetri θ jaoks nihketa, kui $E\hat{\theta} = \theta$. (Lepik, Traat, 2013, lk 22)

Kui üldkogumi maht N pole teada, siis kasutatakse alternatiivset hinnangut:

$$\hat{\bar{Y}}_{\text{alt}} = \frac{\hat{t}}{\hat{N}}. \quad (2)$$

Alternatiivse hinnangu dispersioon ja selle hinnangu valemid järelduvad nn suhtehinnangu dispersiooni valemitest:

$$V(\hat{\bar{Y}}_{\text{alt}}) = \frac{1}{N^2} \sum \sum_U \Delta_{ij} \left(\frac{y_i - \bar{Y}}{E(I_i)} \right) \left(\frac{y_j - \bar{Y}}{E(I_j)} \right) \quad (3)$$

$$\hat{V}(\hat{\bar{Y}}_{\text{alt}}) = \frac{1}{\hat{N}^2} \sum \sum_U \check{\Delta}_{ij} \left(\frac{y_i - \hat{\bar{Y}}_{\text{alt}}}{E(I_i)} \right) \left(\frac{y_j - \hat{\bar{Y}}_{\text{alt}}}{E(I_j)} \right), \quad (4)$$

kus $\Delta_{ij} = \text{Cov}(I_i, I_j)$ ja $\check{\Delta}_{ij} = \frac{\Delta_{ij}}{E(I_i I_j)}$.

(Traat, Inno, 1997, lk 80-81)

Üldiselt on hinnang (2) paremate omadustega kui hinnang (1) ja seetõttu on laialdaselt kasutatav praktikas. Paneme tähele, et LJV TTA disaini korral langevad (1) ja (2) kokku. See aga ei kehti PPS disaini korral.

4.1 Keskmise π - ja p -hinnangud

Käesolevas töös kasutatakse üldkogumi keskmise hindamiseks alternatiivset hinnangut (2). Tihti eelistatakse alternatiivset hinnangut ka siis, kui üldkogumi maht on teada, sest see annab täpsema tulemuse kui üldisest hindamisteoreemist tulenev hinnang.

Kogusumma t jaoks kasutame lineaarset nihketa hinnangut $\hat{t} = \sum_U \frac{I_i y_i}{E(I_i)}$. Sõltuvalt disainist (TTA või TGA) kannab see hinnang erinevaid nimetusi.

Definitsioon 6. Kogusumma $t = \sum_U y_i$ π - hinnanguks nimetatakse (π - laiendatud valimiväärtuste) summat $\hat{t} = \sum_U \frac{y_i I_i}{\pi_i}$, kus $i \in s$. (Traat, Inno, 1997, lk 69)

Keskmise alternatiivse hinnangu jaoks on vaja kasutada üldkogumi mahu hinnangut. Nihketa hinnang üldkogumi mahule avaldub järgmiselt:

$$\hat{N} = \sum_U \frac{I_i}{\pi_i},$$

kus $\frac{1}{\pi_i}$ on kaasamistõenäosuste pöördväärtus ehk laiendustegur³.

Seega keskmise alternatiivne hinnang on

$$\hat{Y}_{\text{alt}} = \frac{\hat{t}}{\hat{N}} = \frac{\sum_U \frac{I_i y_i}{\pi_i}}{\sum_U \frac{I_i}{\pi_i}}. \quad (5)$$

LJV TTA jaoks:

$$\hat{N} = \sum_U \frac{I_i}{\pi_i} = \sum_U \frac{I_i}{\frac{1}{N}} = \frac{N}{n} \sum_U I_i = N,$$

kuna fikseeritud mahuga disaini korral $\sum_U I_i \equiv n$.

Tagasipanekuga disainide korral leitakse nn p -hinnang.

Definitsioon 6. Kogusumma $t = \sum_U y_i$ p -hinnanguks nimetatakse p -laiendatud väärtuste $\frac{y_i}{p_i}$ valimikeskmist $\hat{t} = \frac{1}{n} \sum_U \frac{I_i y_i}{p_i}$. (Traat, Inno, 1997, lk 74)

Tagasipanekuga valiku korral sobib üldkogumi mahu hinnanguks kasutada järgmist valemit:

$$\hat{N} = \frac{1}{n} \sum_U \frac{I_i}{p_i}.$$

Seega keskmise alternatiivne p -hinnang avaldub järgmisel kujul:

$$\hat{Y} = \frac{\sum_U \frac{I_i y_i}{p_i}}{\sum_U \frac{I_i}{p_i}}. \quad (6)$$

(Traat, Inno, 1997, lk 78)

Näeme, et hinnang (6) on mittelineaarne. Seetõttu täpseid statistilise omadusi sellel uurida ei saa. Arendades seda Tayloriga ritta punkti (t_y, N) ümbruses ja võttes lineaarse osa, saab näidata, et hinnang (6) on asümptootiliselt nihketa ning hinnangu ligikaudne dispersioon ja dispersiooni hinnang avalduvad valemite (3) ja (4) abil vastavalt. See osa kuulub „Valikuuringute I“ ainekavasse, vaata konspekt lk 37-39.

³ Laiendustegur näitab mitmele üldkogumi objektile valimiväärtus y_i laiendatakse, et üle valimi summeerimisel saada üldkogumi kogusummat. (Traat, Inno, 1997, lk 69)

5. Osakogumid ja keskmise hindamine osakogumites

Valikuuringute teoorias hinnatakse sageli lisaks üldkogumi parameetritele ka eraldi mingit kindlat üldkogumi osa, st osakogumi parameetreid. Alljärgnev tekst osakogumitest põhineb raamatul Traat, Inno (1997) lk 81-84.

Definitsioon 7. Osakogum on üldkogumi alamhulk, mis on fikseeritav tausttunnuse või uuritava tunnuse väärtuste järgi ja mida soovitakse eraldi uurida ning mille kohta soovitakse esitada eraldi hinnanguid. (Traat, Inno, 1997, lk 15)

Osakogumi objektid on samad, mis üldkogumi objektid. Näiteks, kui üldkogumi objektideks on koolid, siis osakogumiteks on mingi kindla tunnuse abil määratud koolid, mitte klassid või õpilased.

Osakogumeid on võimalik moodustada kahel erineval viisi. Üheks võimaluseks on esmalt eraldada üldkogumist osakogumid ja seejärel seal soovitud valikudisaini rakendada. Nii satub osakogum nõ üldkogumi rolli ja uuritavate parameetrite hindamiseks saab kasutada tavalisi valikuteooria hindamismehhanisme. Sellist eelnevat jagamist käsitleb kihtvalik. Praktikas aga võib tekkida olukordi, kus üldkogumi eelnev jagamine osakogumiteks on väga tömahukas, liiga kulukas või info puudumise tõttu võimatu. Osakogumite parameetritele hinnangu vajadus võib tekkida ka alles pärast valimi moodustamist. Sellisel juhul saame osakogumi objektid valimist välja võtta ja moodustada neist osavalimi. Osavalimi maht on juhuslik suurus ning võib esineda olukordi, kus see on liiga väike nõutava hinnangu täpsuse jaoks. Näiteks, kui valim koosneb 500 abiturientist ja osakogumiks on need õpilased, kelle kirjandi tulemus oli üle 85 punkti, siis osavalimi suurus pole ette teada ja võib juhtuda, et tuleb liiga väike.

Olgu üldkogum $U = \{1, \dots, i, \dots, N\}$ jagatud D mittelõikuvaks osaks $U_1, \dots, U_d, \dots, U_D$. Osakogumi U_d maht olgu N_d . Siis kehtivad võrdused $U = \bigcup_{d=1}^D U_d$ ja $N = \bigcup_{d=1}^D N_d$. Iga üldkogumi objektiga seotakse osakogumi indikaator:

$$z_i = \begin{cases} 1, & \text{kui } i \in U_d; \\ 0, & \text{muidu.} \end{cases}$$

See indikaator võimaldab üle minna uuritavale tunnusele y' väärtustega $y'_i = y_i z_i$ ning rakendada uuele tunnusele olemasolevaid teoreetilisi tulemusi.

5.1 LJV TTA korral keskmise hindamine osakogumites

Osakogumi keskvaartuse $\bar{Y}_d = \frac{t_d}{N_d}$ hindamiseks konstrueerime alternatiivse hinnangu valemi

(2) järgi. Selle jaoks on vaja osakogumi kogusumma hinnangut ning osakogumi mahu hinnangut. Osakogumi kogusumma avaldub üldkogumi kogusumma ja osakogumi indikaatori kaudu järgmiselt:

$$t_d = \sum_{U_d} y_i = \sum_U y_i z_i = \sum_U y_i',$$

kus $y_i' = y_i z_i$.

Kogusumma hinnang LJV TTA korral avaldub järgmiselt:

$$\hat{t}_d = N\bar{y}' = \frac{N}{n} \sum_s y_i' = \frac{N}{n} \sum_{s_d} y_i.$$

Osakogumi mahu $N_d = t_z = \sum_U z_i$ hinnang LJV TTA korral avaldub järgmisel kujul:

$$\hat{N}_d = \hat{t}_z = N\bar{z} = Np_d,$$

kus p_d on valimi s_d osakaal.

Seega alternatiivne hinnang osakogumi keskmisele LJV TTA korral avaldub järgmiselt:

$$\hat{\bar{Y}}_{d,alt} = \frac{\hat{t}_d}{\hat{N}_d} = \frac{\frac{N}{n} \sum_{s_d} y_i}{\frac{N n_d}{n}} = \frac{1}{n_d} \sum_{s_d} y_i = \bar{y}_d.$$

Näeme, et keskmise hinnang osakogumis on lihtsalt valimi keskmine selles osakogumis.

Alternatiivse hinnangu osakogumi keskmise dispersioon avaldub:

$$V(\hat{\bar{Y}}_{d,alt}) = \frac{1}{n_d} (1 - f_d) S_{y_{U_d}}^2$$

ning dispersiooni hinnang

$$\hat{V}(\hat{\bar{Y}}_{d,alt}) = \frac{1}{n_d} (1 - f_d) S_{y_{s_d}}^2,$$

$$\text{kus } f_d = \frac{n_d}{N_d} \text{ - valikusuhe,}$$

$$S_{y_{U_d}}^2 = \frac{1}{N_d - 1} \sum_{U_d} (y_i - \bar{y}_d)^2 \text{ - tunnuse } y \text{ osakogumi dispersioon}$$

$$S_{y_{s_d}}^2 = \frac{1}{n_d - 1} \sum_{s_d} (y_i - \bar{y}_d)^2 \text{ - tunnuse } y \text{ osavalimi dispersioon}$$

(Traat, Inno, 1997, lk 100).

Kui N_d pole teada, siis pole ka teada valikusuhe f_d . Suhte f_d hinnanguks sobib kogu valimi valikusuhe f , sest LJV korral kehtib $E(f_d) = f$. (Traat, Inno, 1997, lk 100)

5.2 PPS tagasipanekuga valiku korral keskmise hindamine osakogumites

Analoogselt eelmisele peatükile konstrueerimine osakogumi keskvaartuse $\bar{Y}_d = \frac{t_d}{N_d}$ hindamiseks alternatiivse hinnangu valemi (2) järgi. Selle jaoks on vaja osakogumi mahu ja osakogumi kogusumma hinnanguid. Osakogumi mahu saab kirja panna indikaatortunnuse abil üldkogumi summana järgmiselt:

$$N_d = \sum_U z_i.$$

Rakendades üldist hindamisteoreemi üldkogumi summade hindamiseks, saame PPS TGA valiku korral osakogumi mahu hinnanguks:

$$\hat{N}_d = \sum_U \omega_i z_i = \sum_s \omega_i z_i = \sum_{s_d} \omega_i = \sum_{s_d} \frac{I_i}{E(I_i)} = \sum_{s_d} \frac{I_i}{np_i}.$$

Näeme, et hinnatud osakogumi maht on kaalude summa üle osavalimi s_d .

Nagu eelnevalt mainitud, siis osakogumi kogusumma avaldub $t_d = \sum_U y_i'$, kus $y_i' = y_i z_i$. Selle summa nihketa hinnang on PPS TGA valiku korral järgmine:

$$\hat{t}_d = \sum_U \omega_i y_i' = \sum_s \omega_i y_i' = \sum_{s_d} \frac{I_i y_i}{np_i}.$$

Järelikult alternatiivne hinnang osakogumi keskmisele PPS TGA valiku korral avaldub:

$$\hat{\hat{Y}}_{d,alt} = \frac{\hat{t}_d}{\hat{N}_d} = \frac{\sum_{s_d} \frac{I_i y_i}{np_i}}{\sum_{s_d} \frac{I_i}{np_i}}.$$

Dispersioon ja selle hinnang tuleb valemist (3) ja (4)

$$V(\hat{\hat{Y}}_{d,alt}) = \frac{1}{N_d^2} \sum \sum_{U_d} \Delta_{ij} \left(\frac{y_i - \bar{Y}_d}{np_i} \right) \left(\frac{y_j - \bar{Y}_d}{np_j} \right)$$

$$\text{kus } \Delta_{ij} = -np_i p_j, \text{ kui } i \neq j$$

$$\Delta_{ii} = np_i(1 - p_i)$$

$$\hat{V}(\hat{\hat{Y}}_{d,alt}) = \frac{1}{\hat{N}_d^2} \sum \sum_{s_d} \Delta_{ij} \left(\frac{y_i - \hat{\hat{Y}}_{d,alt}}{np_i} \right) \left(\frac{y_j - \hat{\hat{Y}}_{d,alt}}{np_j} \right),$$

$$\text{kus } \check{\Delta}_{ij} = \frac{\Delta_{ij}}{E(I_i I_j)}$$

$$\Delta_{ij} = -np_i p_j,$$

$$E(I_i I_j) = n(n-1)p_i p_j$$

$$\Delta_{ii} = np_i(1-p_i)$$

$$E(I_i^2) = np_i(1-p_i+np_i).$$

6. Kitsendustega hindamine

Käesolev peatükk põhineb Sõstra (2007) tööil.

Praktikas esineb olukordi, kus samu üldkogumi parameetreid hinnatakse mitmes erinevas uuringus. Uuritavate parameetrite vahel kehtivad teatud seosed. Statistika tarbijad soovivad, et need seosed kehtiksid ka hinnangute vahel. Näiteks, leibkondade eelarve uuringust saadud summaarne hinnang sissetulekule palgatööst võiks võrduda tööjõu uuringust saadud palkade kogusummaga (või selle hinnanguga). Analoogiliselt, samast üldkogumist leitud hinnangud peavad rahuldama mingeid tingimusi. Näiteks, kui uuritavaks tunnuseks on palk, siis osakogumite palkade kogusumma hinnangud peavad kokku andma üldkogumi palga kogusumma või hinnangu. See aga pole alati täidetud hinnangute jaoks:

1) valimi juhuslikkuse tõttu ei pruugi hinnangud osakogumites summeeruda registris pärit teadaolevaks näitajaks.

2) osakogumites rakendatud hinnangufunktsioon pole lineaarne. Viimane väide kehtib ka suhtehinnangute korral osakogumites.

Üks võimalik lahendus sellele probleemile pakuti välja raamatus Knottnerus (2003), mis kannab nimetust üldine kitsendustega hinnang. Üldist kitsendustega hinnangut nimetame edaspidi GR (General Restriction) - hinnanguks.

6.1 GR - hinnang

GR – hinnangufunktsiooni üldise kuju tutvustus pärineb Sõstra (2007) tööst.

Olgu $\theta = (\theta_1, \theta_2, \dots, \theta_k)^T$ uuritavate parameetrite vektor. Tähistame $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k)^T$ - k -mõõtmeline nihketa hinnangute vektori ühest või rohkematest valimitest. Olgu $\text{Cov}(\hat{\theta}) = V$, kus V on mittesingulaarne⁴ kovaratsiooni maatriks. Eeldame, et parameetrivektor rahuldab järgmist lineaarset kitsendust:

$$R\theta = c,$$

kus R on $r \times k$ maatriks astakuga r ja c on r – mõõtmeline konstantide vektor.

⁴ Mittesingulaarne maatriks – ruutmaatriks on mittesingulaarne, siis ja ainult siis, kui tema determinant ei võrdu nulliga. (Lütkepohl, 1996)

Teoreem 1. (Knottnerus, 2003, lk 328-329) Üldine kitsendustega hinnang $\hat{\theta}^{GR}$, mis rahuldab tingimust $R\theta = c$ ja selle hinnangu kovaratsiooni maatriks V^{GR} avalduvad järgmiselt:

$$\hat{\theta}^{GR} = \hat{\theta} + K(c - R\hat{\theta}),$$

$$V^{GR} = \text{Cov}(\hat{\theta}^{GR}) = (I_k - KR)V,$$

kus I_k on ühikmaatriks ja $K = VR^T(RVR^T)^{-1}$.

Sõstra (2007) näitab oma töös (lk 44-45), et kitsendustega hinnang $\hat{\theta}^{GR}$ on efektiivsem kui esialgne hinnang $\hat{\theta}$, $V^{GR} < V$ Löwneri järjestuse⁵ mõttes.

Praktilistes olukordadest ei teata tavaliselt kovaratsiooni maatriksit V , aga teda saab hinnata valimi põhjal. Sõstra töös (lk 45-46) on näidatud, et kasutades kovaratsiooni maatriksi V asemel tema hinnangut \hat{V} , saadakse ikka hinnangufunktsioon, mis rahuldab kitsendusi.

Näide 1. Olgu osakogumi U_d keskmine $\bar{Y}_d = \frac{t_d}{N_d}$, seega osakogumi kogusumma avaldub järgmiselt:

$$t_d = \bar{Y}_d N_d.$$

Teame, et $\sum_{d=1}^D t_d = t$, kus $t = \sum_U y_i$ on üldkogumi kogusumma. Sellisel juhul $R = (1 \ 1 \ \dots \ 1)_{D \times 1}$ ja $c = t$, $\theta = (t_1, t_2, \dots, t_D)^T$. Parameetrite korral kehtib $R\theta = c$.

Hinnangu vektori korral see tingimus aga pole enam täidetud. Olgu $c = t$ (teadaolev üldkogumi summa, mis on teada registrist).

1) Olgu $\hat{t}_d = \sum_{U_d} \frac{I_i y_i}{E(I_i)}$, siis

$$\sum_{d=1}^D \hat{t}_d = \sum_{d=1}^D \sum_{U_d} \frac{I_i y_i}{E(I_i)} = \sum_U \frac{I_i y_i}{E(I_i)} = \hat{t} \neq t.$$

Sellisel juhul summeeruvad osakogumi summa hinnangud üldkogumi summa hinnanguks, aga mitte teadaolevaks kogusummaks.

⁵ Löwneri järjestus $A < B$, maatriksite A ja B korral, tähendab seda, et maatriks $B - A$ on positiivselt määratud. (Lütkepohl, 1996)

2) Panime tähele, et $t_d = \bar{Y}_d N_d$. Kuna osakogumi keskmise \bar{Y}_d hindamiseks kasutatakse alternatiivset hinnangut (2), siis saame ka alternatiivse hinnangu kogusummale:

$$\hat{t}_{d,alt} = \hat{\bar{Y}}_{d,alt} N_d = \frac{\hat{t}_d}{N_d} N_d. \quad (7)$$

See hinnang on väga levinud praktikas oma väiksema varieeruvuse tõttu võrreldes tavalise lineaarse hinnanguga \hat{t}_d . Hinnang (7) pole aga aditiivne, st

$$\begin{aligned} \sum_{d=1}^D \hat{t}_{d,alt} &\neq t \text{ ega} \\ \sum_{d=1}^D \hat{t}_{d,alt} &\neq \hat{t}. \end{aligned}$$

Antud töös rakendatakse praktilises osas kitsendsutega hinnangut PPS TGA disaini kogusumma hinnangutele tervishoiuteenuse osutaja liigi järgi moodustatud kihtides nii, et oleks täidetud järgmine kitsendus:

$$\hat{t}_1 + \hat{t}_2 + \hat{t}_3 + \hat{t}_4 + \hat{t}_5 + \hat{t}_6 = t - t_{hambaravi},$$

kus \hat{t}_i on i 'nda tervishoiuteenuse osutaja liigi põhitunnipalga kogusumma hinnang, t on üldkogumist pärit põhitunnipalga kogusumma ja $t_{hambaravi}$ on hambaravi põhitunnipalga kogusumma.

7. Hinnangute täpsusnäitajad

Hinnangute täpsusnäitajate valemid ja kirjeldused põhinevad raamatul Traat, Inno (1997) lk 60 - 61.

Valimite abil saadud statistilised hinnangud ei lange kunagi päris täpselt kokku vastavate üldkogumist saadud parameetritega. Olgu hinnatav parameeter θ . Hinnangu viga realiseerunud valimi s korral on hinnangu ja tegeliku parameetri vahe:

$$\hat{\theta}(s) - \theta.$$

Hinnangu viga koosneb kahest komponendist – valikuviga ja muu viga. Valikuviga on põhjustatud sellest, et üldkogumi asemel vaadatakse valimit ning muu viga tuleneb praktiliselt valikuuringu keskkonnast ja vahenditest. Muu vea täpset suurust ei saa hinnata, on võimalik vaid öelda näiteks, kas puudujäägid valikuuringu tegemisel põhjustavad parameetrite üle- või alahindamist.

Valikuviga on võimalik tõenäosusliku valiku korral arvuliselt hinnata. Kui kasutada parameetrite hindamisel nihketa hinnanguid, siis on valikuviga põhjustatud statistiku varieeruvusest üle kõikvõimalike valimite ning seda mõõdetakse standardhälbe mitmesuguste modifikatsioonidega, näiteks suhtelise veana.

Suhteline viga tähistatakse CV , mis tuleneb ingliskeelsest väljendist *coefficient of variation*, ning leitakse järgmise valemi abil:

$$CV(\hat{\theta}) = \frac{\sqrt{\hat{V}\hat{\theta}(s)}}{\hat{\theta}(s)},$$

kus $\sqrt{\hat{V}\hat{\theta}(s)}$ on statistiku standardviga ehk statistiku standardhälbe hinnang.

Suhteline viga näitab, kui suure osa moodustab hinnangu standardviga hinnangust endast. Suhtelist viga väljendatakse protsentides ning teda peetakse operatiivseks vahendiks kiire ülevaate saamiseks hinnangute täpsusest.

Tihti leitakse lisaks parameetri hinnangule ka tema usaldusvahemik. Usaldusvahemik katab parameetri väärtused, mingi teadaoleva tõenäosusega ehk usaldusnivooga. Usaldusvahemik määratakse usalduspiiridega.

Kui statistik on ligikaudu normaaljaotusega ja nihketa, siis parameetri θ usalduspiirid usaldusnivool $1-\alpha$ on:

$$\hat{\theta}(s) \pm \lambda_{\alpha/2} \sqrt{\hat{V}\hat{\theta}(s)},$$

kusjuures alumise usalduspiiri määrab vahe ja ülemise summa.

Usalduspiiride arvutamisel kasutatakse $\lambda_{\alpha/2}$ on standardnormaaljaotuse täiendkvantiil. Tihti kasutatakse α väärtus on 0,05, sellisel juhul $\lambda_{\alpha/2} = 1,96$.

Käesoleva töö praktilises osas leitakse igas osakogumis saadud hinnangule suhteline viga ning usaldusvahemik.

8. Simuleerimisülesanne

Käesolevas tööosas võrreldakse kirjeldatud valiku meetodeid omavahel. Kasutatakse nii tagasipanekuta kui tagasipanekuga disaine – lihtne juhuslik valik, suurusega võrdeline tõenäosusega valik, lihtne juhuslik kihtvalik ja suurusega võrdeline tõenäosusega kihtvalik. Lisaks erinevatele disainidele rakendatakse kitsendustega hinnangut. Hinnangud leitakse erinevates osakogumites. Simulatsiooni abil uuritakse hinnangute keskmist tulemust ja täpsust osakogumites. Täpsuse näitajaks kasutatakse suhtelist viga. Osakogumites saadud hinnangutele leitakse ka 95-% usaldusvahemikud.

8.1 Andmestiku kirjeldus

Käesolevas töös kasutatavad üldkogumi andmed pärinevad Tervise Arengu Instituudilt (2013 aasta seisuga). Iga aasta märtsikuus on tervishoiuteenust osutavad asutused, isegi kui nende põhitegevus ei ole tervishoid, kohustatud täitma statistilist aruannet „Tervishoiutöötajate tunnipalk“. Aruannet saab täita ja esitada sotsiaalministeeriumi kodulehel, kus on selleks loodud veebipõhine tervishoiualase statistika kogumise keskkond A-veeb (www.sm.ee/aveeb). Aruandes kajastatakse nende arstide, hambaarstide, arst-residentide, õdede, ämmaemandate, hooldajate, proviisorite, farmatseutide ja teiste tervishoiutöötajate andmed, kes märtsikuu jooksul on olnud tööandja koosseisu nimekirjas. Aruannet täidetakse ankeedi vormis, tühi ankeet on esitatud lisas 1 (tabel 20). Antud töös käsitletavat andmed sisaldasid ainult arstide, õdede ja hooldajate andmeid. Eelmisel aastal (2013) jättis 20 asutust oma andmed esitamata või esitas nullaruande ning neljal asutusel polnud kohustust oma andmeid esitada.

Huvi pakkuvateks tunnusteks on eelkõige palgaga seotud andmed. TAI-lt saadud andmestikus oli 16878 töötajat ja 38 tunnust. Üldkogumi moodustavad töötajad, kelle põhitunnipalk võeti TAI andmete järgi arvestusse. Järele jäi 747 meditsiiniga seotud asutust, kus 16661 isikul oli põhipalka arvestatud.

Antud töös on kasutatud andmestikust 12 tunnust:

- arvestus – binaarne tunnus. 1 – ei ole põhitunnipalga arvestuses sees, 2 – on põhitunnipalga arvestuses sees. (v2lja_kogutunnipalk)
- jrk – töötaja järjekorra number (@2)
- sugu – binaarne tunnus. 3 – mees, 4 – naine. (@3)
- vanus – vanus (@4)

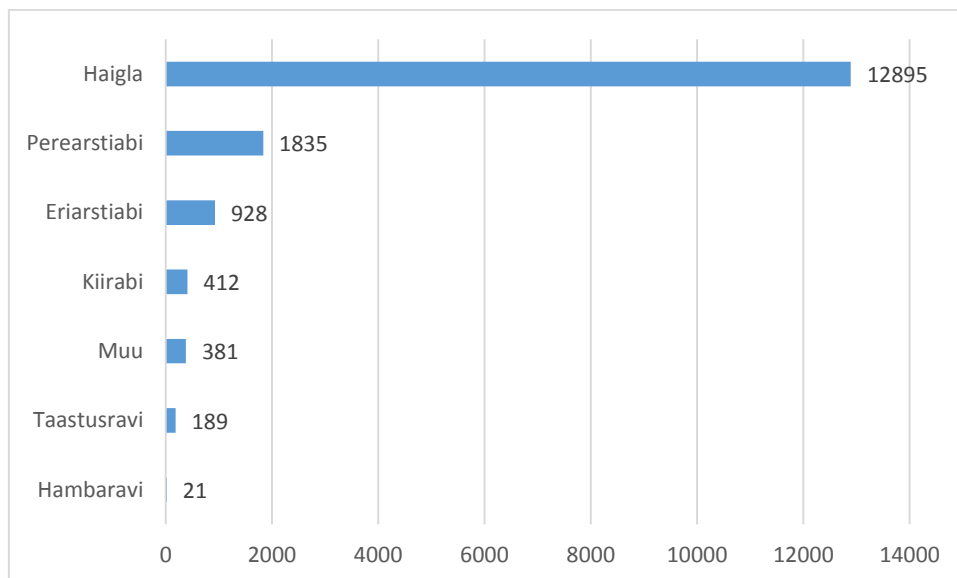
- haridus – haridustase (@8)
- töötunnid – märtsikuu töötunnid (@10)
- ületunnid – märtsikuu töötundide hulgas ületunnid (@11)
- valvetunnid – märtsikuu töötundide hulgas valvetunnid (@12)
- Põhipalk – põhipalk e töölepingus või õigusaktis kindlaks määratud tunni-, päeva-, nädala- või kuupalga määra alusel arvestatud brutopalk märtsikuus (@14)
- Teenus – tervishoiuteenuse osutaja (TTO) liik (nt. haigla, perearstiabi, eriarstiabi, hambaravi, kiirabi, taastusravi, muu) (Teenus_uus)
- Haigla liik – haigla liigid (nt. piirkonlik haigla, keskhaigla, üldhaigla, taastusravihaigla, hooldushaigla, erihaigla, kohalik haigla) (Haigla_liik)
- Ametigrupp – arstid, õed ja hooldajad (Ametigrupp)

Üldkogumi moodustamiseks kasutati tunnust arvestus, ehk need kellel võrdus vastav rida kahega jäid üldkogunnisse sisse. Kuna tunnuses töötaja järjekorra number esinesid mõningased vead, pidime tunnuse haigla suuruse jaoks moodustama igale inimesele unikaalse koodi, selleks kasutasime tunnuseid jrk, sugu, vanus ja haridus. Tunnuseid töötunnid, ületunnid, valvetunnid ning põhipalk kasutati uuritava tunnuse põhitunnipalk moodustamiseks:

$$\text{Tunnipalk} = \frac{\text{põhipalk}}{\text{töötunnid} - \text{ületunnid} - \text{valvetunnid}}.$$

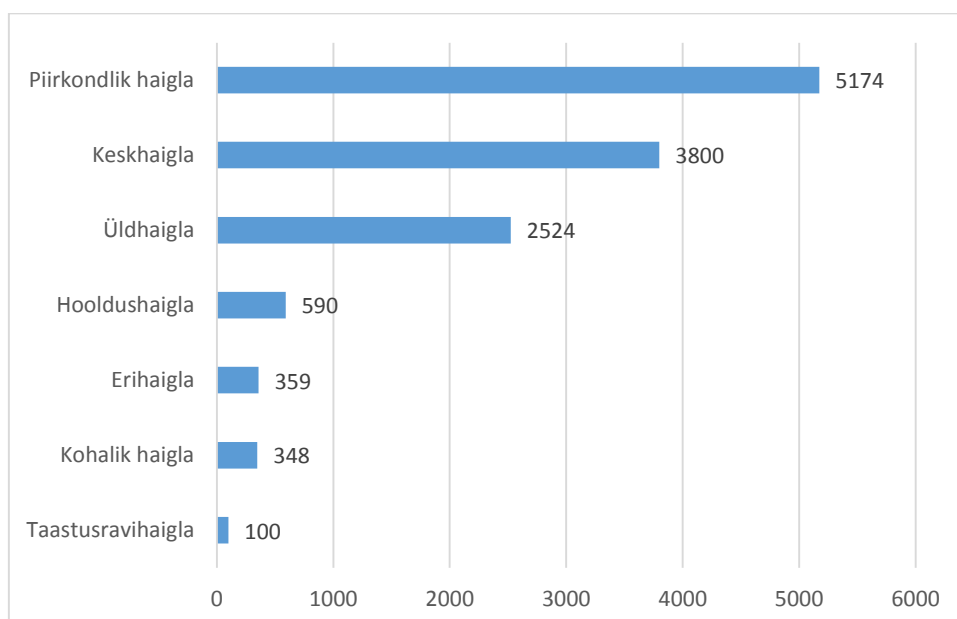
Osakogumite moodustamiseks on kaks jaotust. Esiteks moodustatakse osakogumid tervishoiuteenuse osutaja liigi järgi, lisaks vaadatakse eraldi haigla osakogumis erinevate haigla liikide hinnanguid. Teiseks moodustatakse osakogumid ametigruppide kaupa – arstid, õed ja hooldajad.

Andmestikus on kokku seitse tervishoiuteenuse liiki. Enim, 77% meditsiinitöötajad töötavad haiglates ning kõige vähem, 0,1% on hambaraviteenuse töötajaid. Täpsemat töötajate arvu ning nende osakaalu igas tervishoiuteenuse liigis kajastab joonis 1 ja lisa 2 tabel 17.



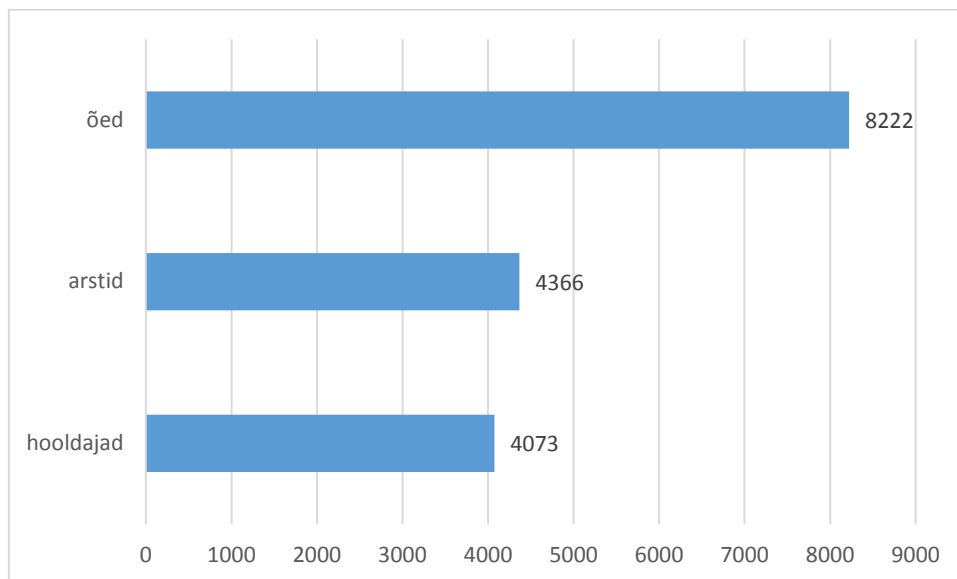
Joonis 1. Töötajate jagunemine TTO liigi järgi

Tervishoiuteenuse osutaja liik haigla jaguneb veel omakorda seitsmeks erinevaks liigiks. Enim, 40% haiglates töötavatest meditsiinitöötajatest töötavad piirkondlikutes haiglates, keskhaiglates töötab 29% ning üldhaiglates 19%. (joonis 2 ja lisa 2 tabel 18)



Joonis 2. Töötajate jagunemine erinevates haigla liikides

Käesoleva andmestiku järgi on ligi pooled tervishoiuteenust pakkuvate asutuste meditsiinitöötajatest õed, hooldajate ja arstite osakaal on vastavalt 25% ja 26%, vt joonis 3.



Joonis 3. Meditsiinitöötajate arvud erinevates ametigruppides

Tervishoiuteenuse osutaja liigi järgi saavad keskmiselt kõige kõrgemat põhitunnitasu eriarstiabis töötavad meditsiinitöötajad ning kõige väiksemat taastusravis töötajad. Täpsed andmed on toodu tabelis 1.

Tabel 1. TTO liigi järgi keskmine tunnipalk

TTO liik	Haigla	Perearstiabi	Eriarstiabi	Hambaravi	Kiirabi	Taastusravi	Muu
Keskmine tunnipalk	5,11	6,84	8,95	5,17	5,50	4,80	5,99

Lisaks vaatasime ka eraldi haigla liigi järgi tervishoiutöötajate keskmist tunnitasu. Enim teenivad meditsiinitöötajad erihaiglates (tabel 2).

Tabel 2. Haigla liigi järgi keskmine tunnipalk

Haigla liik	Piirkondlik	Keskhaigla	Üldhaigla	Taastusravi- haigla	Hooldus- haigla	Erihaigla	Kohalik haigla
Keskmine tunnipalk	5,21	5,23	4,89	4,23	4,19	6,49	4,40

Tunnus ametigrupp sisaldab arste, õdesid ja hooldajaid. Neist kõrgeim keskmine tunnipalk on arstidel ja madalaim hooldajatel. Arstide keskmine põhitunnipalk ületab õdede põhitunnipalka kahekordselt ning hooldajate oma ligi neljakordselt. (tabel 3)

Tabel 3. Ametigrupi järgi keskmine tunnipalk.

Ametigrupp	Arstid	Õed	Hooldajad
Keskmine tunnipalk	9,86	4,72	2,58

Näeme, et arstide tunnitase on tunduvalt kõrgem, kui hooldajatel ja õdedel, sellest tulenevalt on ka eriarstiabi tunnitase kõrgem, sest seal on 55% meditsiini töötajatest arstid. Kuna andmestikus tunnuse ametigrupp all pole hambaarste eraldi välja toodud, siis järelikult hambaravi keskmine tunnitase peegeldab pigem hambaravis töötavate õdede ja hooldajate keskmist tunnitase. Seda näitab ka töötajate osakaalu jagunemine hambaravi grupis, nimelt 2 arsti, 7 õde ja 12 hooldajat ehk 57% töötajatest moodustavad hooldajad. (lisa 2 tabel 17)

8.2 Valimi võtmise ja parameetrite hindamine

Käesolevas töös leitakse simuleerimise abil keskmine tunnipalga hinnang huvipakkuvates osakogumites. Hinnang leitakse LJV TTA, PPS TGA disainide korral. Lisaks teostatakse ka kihtvalik, ühel juhul võetakse igast kihist lihtne juhuslik valik tagasipanekuta ning teisel juhul suurusega võrdeline tõenäosuslik valik tagasipanekuga. Rakendatakse ka kitsendustega hinnangud tunnitase summale.

8.2.1 Keskmise põhitunnipalga hinnang LJV TTA korral

LJV TTA on kõige lihtsam disain teostamise ja analüüsimise mõttes. Samas ei pruugi see anda kõige paremat tulemust väikestes osakogumites. Antud töös on LJV TTA disaini kasutatud võrdlemiseks teise disainiga (PPS TGA).

Esmalt simuleeriti lihtsat juhuslikku valiku tagasipanekuta disaini kasutades 1 000 valimit. Valimi suuruseks võeti 30% üldkogumi mahust ehk 4 999 inimest. Seejärel moodustati igas valimis osakogumid. Osakogumid moodustati tunnuse tervishoiuteenuse osutaja liigi järgi. Tekkis seitse osavalimit – haigla, perearstiabi, eriarstiabi, hambaravi, kiirabi, taastusravi ja muu. Leides osavalimites tunnitase keskmised saadi hinnangud osakogumitele. Lisaks leiti ka iga hinnangu jaoks suhteline viga protsentides ja vahemikhinnangud. Saadud tulemused on toodud tabelis 4.

Tabel 4. Tunnitase hinnangud ja täpsusnäitajad TTO liigi järgi.

TTO liik	Keskmine tunnipalga hinnang	Dispersioon	Suhteline viga	Usaldusvahemiku alumine piir	Usaldusvahemiku ülemine piir
Haigla	5,11	0,003	1,0%	5,01	5,21

TTO liik	Keskmine tunnipalga hinnang	Dispersioon	Suhteline viga	Usaldusvahemiku alumine piir	Usaldusvahemiku ülemine piir
Eriarstiabi	8,94	0,14	4,2%	8,20	9,69
Perearstiabi	6,83	0,01	1,3%	6,65	7,01
Hambaravi	5,14	1,42	23,2%	2,81	7,48
Kiirabi	5,50	0,01	2,1%	5,28	5,72
Taastusravi	4,80	0,08	6,0%	4,24	5,37
Muu	5,99	0,04	3,4%	5,60	6,39

Näeme, et hambaravi osakogumis on suhteline viga üpris suur, 23,2%, st hinnangu standardviga on peaaegu veerand hinnangust endast. Seda peetakse väga halvaks täpsuseks. Samuti näeme, et 95% tõenäosusega jääb hambaravis töötavate meditsiinitöötajate tegelik tunnitasu 2,81 ja 7,48 euro vahele, mis on suur erinevus. Teistes osakogumites saadud hinnangud on hea täpsusega.

Leidsime hinnangud ka eraldi haigla liikide järgi. Kõige ebatäpsem hinnang tuli taastusravi haiglatele, näeme usaldusvahemikust, et 95% tõenäosusega jääb LJV TTA kasutades keskmise tunnipalga hinnang vahemikku 3,5 - 4,9 eurot, mis on üpris laia usaldusvahemik. Tunnitasi kontekstis on suur erinevus, kas saadakse 1,4 eurot tunnis rohkem või vähem. See mõjutab väga tugevalt kuupalka. Parim saadud täpsustest on piirkondlikutel haiglatel, suhteline viga on vaid 1,1%. Samuti üsna väikse suhtelise veaga, täpsemal 1,7% on üldhaiglatele saadud hinnang. (tabel 5)

Tabel 5. Tunnitasi hinnangud ja täpsusnäitajad haigla liigi järgi.

Haigla liik	Keskmine tunnipalga hinnang	Dispersioon	Suhteline viga	Usaldusvahemiku alumine piir	Usaldusvahemiku ülemine piir
Piirkondlik	5,21	0,004	1,1%	5,10	5,33
Keskhaigla	5,22	0,02	2,5%	4,96	5,48
Üldhaigla	4,89	0,01	1,7%	4,73	5,05
Taastusravi	4,21	0,13	8,5%	3,51	4,91
Hooldus	4,18	0,06	6,0%	3,69	4,67
Erihaigla	6,49	0,08	4,5%	5,92	7,06
Kohalik	4,39	0,04	4,5%	4,00	4,77

Ametigrupi järgi moodustatud osakogumites tulid hinnangute täpsused väga head. Õdede ja hooldajate tunnitasi suhtelised vead olid vastavalt 0,4% ja 0,5%. 95% tõenäosusega jääb õdede tunnitasi vahemikku 4,59 kuni 4,84 eurot. Ülevaate keskmise tunnipalga hinnangutest ja täpsusnäitajatest ametigruppides annab tabel 6.

Tabel 6. Tunnitasku hinnangud ja täpsusnäitajad erinevates ametigruppides LJV TTA

Ametigrupp	Keskmine tunnipalga hinnang	Dispersioon	Suhteline viga	Usaldusvahemiku alumine piir	Usaldusvahemiku ülemine piir
Arstid	9,85	0,0178	1,4%	9,62	10,08
Õed	4,72	0,0004	0,4%	4,59	4,84
Hooldajad	2,58	0,0002	0,5%	2,44	2,72

Enamus hinnangute ja täpsusnäitajatega võib rahule jääda, v.a hambaravi hinnang. Hambaravi hinnangu suhteline viga oli üle 20%, mis näitab tulemuse halba täpsust.

Saadud tulemuste SAS'i koodiga saab tutvuda lisas 3 1.kood: LJV TTA hinnangute leidmine.

8.2.2 Keskmise põhitunnipalga hinnang PPS TGA korral

Tervise Arengu Instituudi seisukohalt mõjutavad suuremad tervishoiuasutused keskmist palka rohkem võrreldes väiksemate (meditsiinitöötajate arvu poolest) asutustega. Seetõttu soovisid nad, et valimisse sattuksid pigem suuremate haiglate töötajad. Seega otsustati läbi viia haigla suurusega võrdeline tõenäosuslik valik.

Tunnitasku hinnangu saamiseks simuleerisime samuti 1 000 valimit PPS tagasipanekuga disaini abil. Valimi maht jäi samaks ehk 4 999 inimest. Tausttunnusena kasutati haigla suurust. Tunnus haigla suurus leiti vastava haigla meditsiinitöötajate loendamise teel. Paraku olid mõningased vead tunnuses töötaja järjekorra number, seega pidime igale töötajale moodustama unikaalse koodi. Näiteks esines olukord, kus töötaja järjekorra number oli haiglates sama, aga tema sugu, vanus ning haridustase olid erinevad. Seega kasutasimegi neid tunnuseid unikaalse koodi loomiseks. Iga töötaja sai endale koodi, mis peegeldas tema endist järjekorra numbrit, sugu, vanust ja haridustaset. Igale haiglale leiti töötajate arv järgmiselt: loendati unikaalse koodi veergu, kui mõnel real kood kattus varajasemaga, siis seda uuesti ei loeta. Seega saime teada mitu inimest tegelikult töötab vastavas haiglas.

Suurusega võrdelise tõenäosusega tagasipanekuga disaini korral tulid hinnangute täpsused paljudes osakogumites tunduvalt halvemad kui LJV TTA korral. Hinnangud ja nende täpsusnäitajad erinevates osakogumites on välja toodud tabelites 7, 8 ja 9.

Tabel 7. Tunnitasku hinnangud ja täpsusnäitajad TTO liigi järgi PPS TGA korral

TTO liik	Keskmine tunnipalga hinnang	Dispersioon	Suhteline viga	Usaldusvahemiku alumine piir	Usaldusvahemiku ülemine piir
Haigla	5,11	0,02	2,5%	4,86	5,36
Eriarstiabi	8,56	16,12	46,9%	0,69	16,43
Perearstiabi	6,84	2,33	22,3%	3,84	9,83
Hambaravi	4,92	8,49	59,2%	0,00	10,63
Kiirabi	5,51	0,20	8,2%	4,63	6,39
Taastusravi	4,54	4,32	45,8%	0,47	8,62
Muu	5,98	1,70	21,8%	3,42	8,54

Tabel 8. Tunnitasku hinnangud ja täpsusnäitajad haigla liikide järgi PPS TGA

Haigla liik	Keskmine tunnipalga hinnang	Dispersioon	Suhteline viga	Usaldusvahemiku alumine piir	Usaldusvahemiku ülemine piir
Piirkondlik	5,21	0,003	1,0%	5,10	5,31
Keskhaigla	5,23	0,03	3,1%	4,92	5,55
Üldhaigla	4,89	0,06	4,8%	4,43	5,35
Taastusravi	4,12	2,99	41,9%	0,74	7,51
Hooldus	4,17	4,16	48,9%	0,17	8,17
Erihaigla	6,39	2,22	23,3%	3,47	9,31
Kohalik	4,37	0,66	18,6%	2,77	5,97

Tabel 9. Tunnitasku hinnangud ja täpsusnäitajad ametigruppide kaupa PPS TGA

Ametigrupp	Keskmine tunnipalga hinnang	Dispersioon	Suhteline viga	Usaldusvahemiku alumine piir	Usaldusvahemiku ülemine piir
Arstid	9,84	0,99	10,1%	7,89	11,79
Õed	4,71	0,04	4,1%	4,33	5,09
Hooldajad	2,58	0,00	2,5%	2,45	2,70

Saadud hinnangute põhjal võib järeldada, et PPS TGA rakendatuna tervele üldkogumile ei anna häid hinnanguid soovitud osakogumites. Seega peaks esmalt üldkogumi jagama kihtideks, mis on määratud osakogumitega ja seejärel võtta valimid igast osakogumist eraldi.

8.2.3 Keskmine põhitunnipalga hinnang lihtsa juhusliku kihtvaliku korral

Kihtideks jagamine toimus tunnuse tervishoiuteenuse osutaja liigi järgi. Tervise Arengu Instituudil on võimalik tervishoiutöötajate kohta vajaminevat infot saada näiteks

terviseametilt. Lisaks teostavad nad ka ise tervishoiutöötajate kõikset uuringut. Seega moodustati kihid inimeste põhiselt.

Kihtideks jagati võrdelise paigutuse abil, st vastavate kihtide osakaalud valimis ja üldkogumis on võrdsed ehk

$$n_h = n \frac{N_h}{N}.$$

Valimi maht on endiselt 30% üldkogumist ehk 4 999 inimest. Seega kihtide mahud tulevad järgmised:

- Haigla – 3 869
- Perearstiabi – 550
- Eriarstiabi – 278
- Hambaravi – 7
- Kiirabi – 124
- Taastusravi – 57
- Muu – 114

Igas kihis simuleeriti 1 000 valimit LJV TTA disaini kasutades. Leiti põhitunnipalga hinnang igas kihis ja täpsusnäitajad. Samuti leiti põhitunnipalga hinnang ja soovitud täpsusnäitajad erinevates haigla liikides ning ametigrupi järgi moodustatud osakogumites. Saadud tulemused on esitatud tabelites 10, 11 ja 12.

Tabel 10. Põhitunnipalga hinnangud ja täpsusnäitajad TTO kihtides LJV TTA korral

TTO liik	Keskmine tunnipalga hinnang	Dispersioon	Suhteline viga	Usaldusvahemiku alumine piir	Usaldusvahemiku ülemine piir
Haigla	5,11	0,002	1,0%	5,01	5,21
Eriarstiabi	8,94	0,14	4,2%	8,21	9,67
Perearstiabi	6,83	0,01	1,3%	6,66	7,01
Hambaravi	5,11	0,96	19,2%	3,19	7,04
Kiirabi	5,50	0,01	2,2%	5,26	5,74
Taastusravi	4,79	0,08	6,1%	4,22	5,36
Muu	5,99	0,04	3,3%	5,60	6,38

Tabel 11. Põhitunnipalga hinnangud ja täpsusnäitajad erinevates haigla liikides LJV TTA kihtvaliku korral

Haigla liik	Keskmine tunnipalga hinnang	Dispersioon	Suhteline viga	Usaldusvahemiku alumine piir	Usaldusvahemiku ülemine piir
Piirkondlik	5,21	0,003	1,1%	5,09	5,32
Keskhaigla	5,22	0,02	2,5%	4,97	5,47
Üldhaigla	4,89	0,01	1,7%	4,72	5,05
Taastusravi	4,24	0,13	8,4%	3,54	4,93
Hooldus	4,20	0,06	6,0%	3,71	4,69
Erihaigla	6,50	0,08	4,5%	5,93	7,06
Kohalik	4,40	0,04	4,3%	4,02	4,77

Tabel 12. Põhitunnipalga hinnangud ja täpsusnäitajad ametigrupiti LJV TTA kihtvaliku korral

Ametigrupp	Keskmine tunnipalga hinnang	Dispersioon	Suhteline viga	Usaldusvahemiku alumine piir	Usaldusvahemiku ülemine piir
Arstid	9,75	0,02	1,6%	9,44	10,05
Õed	4,60	0,0005	0,5%	4,56	4,64
Hooldajad	2,56	0,0002	0,6%	2,54	2,59

Eelnev üldkogumi kihistamine parandas dispersioone ja suhtelist viga. Hambaravi kihi suhteline viga on 19,2%, mis on küll parem kui LJV TTA rakendatult üldkogumile korral saadud tulemus, aga siiski veel üpris suur. Ülejäänud kihtides ja osakogumites saadud hinnangud on head.

8.2.4 Keskmine põhitunnipalga hinnang PPS TGA kihtvaliku korral

Hinnangute parandamise eesmärgil otsustati taas üldkogum eelnevalt kihtideks jagad, aga seekord võtta valim PPS TGA disaini abil. Kihid moodustati tervishoiuteenuse osutaja liigi järgi analoogselt eelmisele kihistamisele. Tekkis seitse erinevat kihti – haigla, perearstiabi, eriarstiabi, hambaravi, kiirabi, taastusravi ja muu. Kuna hambaravi kihis on vähe objekte (21 inimest), ning võttes sealt veel omakorda valimi PPS TGA disainiga 30%, tuli keskmise põhitunnipalga hinnangul väga suur varieeruvus. Seega otsustati, et hambaravi kiht tuleb võtta kõikselt valimisse. Samuti leiti hinnangud ka haigla kihtides eraldi. Kuna taastusravi haiglate, hooldushaiglate, erihaiglate ja kohalike haiglate kihis tulid halvad hinnangud, otsustati ka need kihid panna kõikselt valimisse. Sooviti, et valimi maht ei ületaks 30% üldkogumi mahust. Seega summeeriti kihtide mahud, mis satuvad valimisse kõikselt.

$$n_{kõikne} = n_{hambaravi} + n_{taastusravi} + n_{hooldush} + n_{erihaigla} + n_{kohalik}$$

$$= 21 + 100 + 590 + 359 + 348$$

Tulemuseks saadi 1 418, mis on 8,5% üldkogumi mahust (16 661-st). Järelikult jääb meil valimi võtmiseks veel 21,5% üldkogumi mahust, milleks on 3 582 objekti. Valim võetakse 6 kihist võrdelise paigutuse abil. Saadud kihtide suurused on järgmised:

- Haigla – 2 701 objekti
- Perearstiabi – 431 objekti
- Eriarstiabi – 218 objekti
- Kiirabi – 97 objekti
- Taastusravi – 45 objekti
- Muu – 90 objekti.

Leitakse 1 000 valimit ning igas valimis põhitunnipalga hinnang. Vaatame tuhande valimi pealt saadud keskmist hinnangut ja nende varieeruvust. Samuti leiame suhtelise vea ning vahemikhinnangud. Saadud hinnangud kihtide kaupa on näha tabelis 13. Paneme tähele, et kõigis kihis on hinnangute suhteline viga alla 15%, mis on meid rahuldav tulemus.

Tabel 13. Põhitunnipalga hinnangud ja täpsusnäitajad TTO kihtides PPS TGA kihtvaliku korral

TTO liik	Keskmine tunnipalga hinnang	Dispersioon	Suhteline viga	Usaldusvahemiku alumine piir	Usaldusvahemiku ülemine piir
Haigla	5,11	0,01	1,9%	4,92	5,30
Eriarstiabi	9,03	1,66	14,3%	6,51	11,55
Perearstiabi	6,84	0,04	3,1%	6,42	7,25
Hambaravi	5,17	0	0	5,17	5,17
Kiirabi	5,50	0,03	3,1%	5,17	5,84
Taastusravi	4,77	0,35	12,5%	3,61	5,93
Muu	6,01	0,29	8,9%	4,96	7,06

Haigla keskmise tunnipalga hinnangu saame kokku kolmest haigla liigi hinnangust ning neljast kihist, mis sattusid kõikselt valimisse. Seega haiglate hinnang leiti järgmise valemi abil:

$$\hat{Y}_{haigla} = \frac{\hat{t}_{haigla}}{N},$$

$$\text{kus } \hat{t}_{haigla} = \hat{t}_1 + \hat{t}_2 + \hat{t}_3 + \sum_{i=4}^7 t_i.$$

Kuna haigla liikide hinnangud on sõltumatud ($\hat{t}_1 \perp \hat{t}_2 \perp \hat{t}_3$), siis dispersiooni hinnang avaldub järgmiselt:

$$\hat{V}(\hat{t}_{haigla}) = \sum_{i=1}^3 \hat{V}(\hat{t}_i).$$

Ning haigla keskmise põhitunnipalga dispersiooni hinnang avaldub kujul:

$$\hat{V}\left(\frac{\hat{Y}_{haigla}}{N}\right) = \hat{V}\left(\frac{\hat{t}_{haigla}}{N}\right) = \frac{1}{N^2} \sum_{i=1}^3 \hat{V}(\hat{t}_i).$$

Tabelis 14 on esitatud hinnangud haigla liikide kaupa. Hinnangud leiti kolmes haigla liigis – piirkondlikud haiglad, keskhaiglad ning üldhaiglad, ülejäänus osakogumites on välja toodud vastav parameeter üldkogumist.

Tabel 14. Põhitunnipalga hinnangud ja täpsusnäitajad erinevates haigla liikides PPS TGA kihtvaliku korral

Haigla liik	Keskmine tunnipalga hinnang	Dispersioon	Suhteline viga	Usaldusvahemiku alumine piir	Usaldusvahemiku ülemine piir
Piirkondlik	5,21	0,006	1,4%	5,06	5,35
Keskhaigla	5,23	0,05	4,1%	4,81	5,64
Üldhaigla	4,89	0,12	7,1%	4,21	5,56
Taastusravi	4,23	0	0	4,23	4,23
Hooldus	4,19	0	0	4,19	4,19
Erihaigla	6,49	0	0	6,49	6,49
Kohalik	4,40	0	0	4,40	4,40

Tervise Arengu Instituut väljastab põhitunnipalgad ka ametigruppide kaupa, leidsime hinnangud ametigruppides. Ametigruppides leiti hinnangud sarnaselt haigla kihi hinnangule.

Tabel 15. Põhitunnipalga keskmise hinnangud ja täpsusnäitajad ametigrupiti PPS TGA kihtvaliku korral

Ametigrupp	Keskmine tunnipalga hinnang	Dispersioon	Suhteline viga	Usaldusvahemiku alumine piir	Usaldusvahemiku ülemine piir
Arstid	10,49	0,12	3,3%	9,81	11,17
Õed	5,05	0,003	1,1%	4,94	5,16
Hooldajad	2,94	0,001	1,0%	2,89	3,00

Kasutades esmalt üldkogumi kihtideks jagamist tervishoiuteenuse osutaja liigi järgi ning seejärel võttes valimi PPS TGA disaini abil tulid keskmise põhitunnipalga hinnangud rahuldavad, kõik suhtelised vead jäid alla 14,5%.

8.2.5 Kitsendustega hinnang

Kitsendustega hinnangut rakendati põhitunnipalga summa hinnangutele PPS TGA kihtvaliku korral. Hinnangud on leitud 6 erinevas kihis, sest hambaravi kiht on kõikne. Tabelis 16 on esitatud igas kihis põhitunnipalga summa tegelik üldkogumist saadud parameeter.

Tabel 16. Põhitunnipalga kogusummad TTO liikide kaupa

TTO liik	Haigla	Eriarsti- abi	Perearsti- abi	Hamba- ravi	Kiirabi	Taastus- ravi	Muu
Tunnipalga kogusumma	65 907,57	8 306,41	12 544,42	108,49	2 265,50	906,65	2 280,44

Tabelis 17 on toodud põhitunnipalga summa hinnangud PPS TGA kihtvaliku korral. Põhitunnipalga summa hinnangute ja täpsusnäitajate koodiga saab tutvuda lisa 7 5.kood:kitsendustega hinnang.

Tabel 17. Põhitunnipalga summa hinnang ja täpsusnäitajad PPS TGA kihtides

TTO liik	Tunnipalga summa hinnang	Dispersioon	Suht. viga	Usaldusvahemiku alumine piir	Usaldusvahemiku ülemine piir
Haigla	65 888,45	1 545 722,68	1,9%	63 451,64	68 325,26
Eriarstiabi	8 379,98	1 428 630,42	14,3%	6 037,28	10 722,68
Perearstiabi	12 542,85	151 310,46	3,1%	11 780,44	13 305,26
Hambaravi	108,49	0	0	108,49	108,49
Kiirabi	2 267,87	5 087,66	3,1%	2 128,07	2 407,67
Taastusravi	901,53	12 610,80	12,5%	681,42	1 121,63
Muu	2 288,23	41 645,17	8,9%	1 888,25	2 688,21

Kitsendust rakendame tunnitasu kogusummadele vastavates kihtides. Olgu θ_d tunnitassude summa kihis U_d , $\theta_d = \sum_{U_d} \text{tunnipalk}_i$. Olgu c põhitunnipalkade kogusumma üldkogumis võetuna registrist. Kuna hambaravi kiht on kõikne, siis kasutame $c - c_1$, kus c_1 on hambaravi kihis põhitunnipalga kogusumma. Olgu $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_6)^T$, tahame, et $\sum_{d=1}^6 \hat{\theta}_d = c - c_1$. Kitsendustega hinnangu leiame järgmise valemi abil:

$$\hat{\theta}^{GR} = \hat{\theta} + K(c - c_1 - R\hat{\theta}),$$

kus $R = (1 \ 1 \ 1 \ 1 \ 1 \ 1)$ ja $K = VR^T(RVR^T)^{-1}$.

Vektori K leidmiseks on vaja esmalt leida kovaratsiooni maatriks V . Selle leiame üle simulatsioonide:

$$V = \begin{pmatrix} V_1 & \dots & Cov(\hat{\theta}_d, \hat{\theta}_g) \\ \vdots & \ddots & \vdots \\ Cov(\hat{\theta}_d, \hat{\theta}_g) & \dots & V_6 \end{pmatrix}, \text{ kus } V_d = V(\hat{\theta}_d).$$

Saadud kovaratsiooni maatriks V on esitatud tabelis 18.

Tabel 18. Kovaratsiooni maatriks

	Perearstiabi	Eriarstiabi	Kiirabi	Taastusravi	Muu	Haigla
Perearstiabi	151 310,46	-897,83	392,43	673,53	1 411,70	-9 286,65
Eriarstiabi	-897,83	1 428 630,42	3 176,19	4 231,64	-17 187,17	-21 299,60
Kiirabi	392,43	3 176,19	5 087,66	177,49	66,71	-607,79
Taastusravi	673,53	4 231,64	177,49	12 610,80	544,37	5 483,18
Muu	1 411,70	-17 187,17	66,71	544,37	41 645,17	2 144,31
Haigla	-9286,65	-21 299,60	-607,79	5 483,18	2 144,31	1 545 722,68

Pärast kitsenduse rakendamist põhitunnipalga hinnangud muutusid täpsemaks. Kõigis kihtides paranesid dispersioonid, nagu teooriapõhjal oodata võis, ning suhtelised vead. Tulemused on toodud tabelis 19.

Tabel 19. Põhitunnipalga summa hinnangud ja täpsusnäitajad pärast kitsenduse rakendamist.

TTO liik	Põhitunnipalga summa hinnang	Dispersioon	Suhteline viga	Usaldusvahemiku alumine piir	Usaldusvahemiku ülemine piir
Haigla	65 860,21	803 833,30	1,4%	64 102,94	67 617,49
Eriarstiabi	8 354,08	804 035,90	10,7%	6 596,58	10 111,57
Perearstiabi	12 540,19	144 707,30	3,0%	11 794,59	13 285,78
Kiirabi	2 267,72	5 065,64	3,1%	2 128,22	2 407,22
Taastusravi	901,09	12 430,62	12,4%	682,56	1 119,61
Muu	2 287,70	41 382,80	8,9%	1 888,98	2 686,42

9. Kokkuvõte

Antud bakalaureusetöö eesmärgiks oli võrrelda suurusega võrdelise tõenäosusega valikut lihtsa juhusliku valikuga ning seda ka kihtvaliku korral. Neid valimi võtmise viise teostati Tervise Arengu Instituudi poolt kogutud 2013 aasta andmetel, et seeläbi soovitada Tervise Arengu Instituudile sobivat valimi võtmise meetodit. Selleks simuleeriti 1 000 valimit igast disainist ning vaadati saadud hinnangute täpsusnäitajaid. Uuriti tervishoiutöötajate põhitunnipalka kahel erineval viisil moodustatud osakogumites. Osakogumid olid tekitatud tervishoiuteenuse osutaja liigi ja ametigrupi põhjal. Haiglad jagunesid veel omakorda liigipõhiselt seitsmeks.

Tervise Arengu Instituudi poolne soov oli, et valimisse satuksid pigem suuremate haiglate töötajad. Seega võrreldi haigla suurusega võrdelise tõenäosusega valikut lihtsa juhusliku valikuga. Saadud tulemuste põhjal võib kindlalt öelda, et paremad hinnangud andis lihtne juhuslik valik. Hinnangud oli täpsemad nii tervishoiuteenuse osutaja liigi järgi tekitatud osakogumites, haigla liikide osakogumites, kui ka erinevates ametigruppides. Ametigruppide osakogumites LJV TTA disaini korral olid märkimisväärselt täpsed hinnangud, üheski (arstid, õed, hooldajad) osakogumis ei ületanud suhteline viga 1,6% ning õdede osakogumis oli vastav näitaja lausa 0,4% ja hooldajatel 0,5%. Samas polnud hambaravi osakogumile saadud keskmise põhitunnipalga hinnang, 5,14 eurot, piisavalt täpne, selle suhteline viga ületas 20%, mida peetakse valikuuringute teoorias halvaks täpsuseks.

Täpsuse parandamiseks otsustati esmalt üldkogum jagada kihtideks tervishoiuteenuse osutaja liigi järgi ning seejärel võtta igast kihist valim. Seekord võrreldi suurusega võrdelist tõenäosusega kihtvalikut lihtsa juhusliku kihtvalikuga. Mõlema disaini korral tulid erinevates ametigruppides saadud hinnangud alla 20%. Lihtsa juhusliku kihtvaliku korral oli taaskord halvim hambaravi kihi keskmine põhitunnipalga hinnang, mille suhteline viga oli 19,2%. Ülejäänud hinnangute täpsusnäitajad oli head. PPS TGA kihtvaliku korral võeti hambaravi kõikselt valimisse, samuti võeti teatud haigla liigid (taastusravihaiglad, hooldushaiglad, erihaiglad ja kohalikud haiglad) kõikselt valimisse. Selle disaini korral oli halvim suhteline viga (14,3%) eriarstiabi kihis. Enamus kihtides ja osakogumites olid hinnangute täpsused paremad lihtsat juhuslikku kihtvalikut kasutades.

Sooviti veelgi parandada PPS TGA kihtvaliku korral saadud hinnangute täpsust. Selleks rakendati PPS TGA kihtvalikult saadud põhitunnipalga kogusummadele kitsendustega

hinnangut. Pärast kitsenduste rakendamist paranesid hinnangute dispersioonid ja ka suhtelised vead, kuid jäid siiski lihtsa juhusliku kihtvaliku täpsusnäitajatest halvemaks.

Kokkuvõtteks võib öelda, et käesolev töö täitis oma eesmärgi. Antud töö tulemustest lähtudes võib soovitada Tervise Arengu Instituudil kasutada pigem lihtsat juhuslikku kihtvalikut. Lihtsa juhusliku kihtvaliku korral soovitaks võtta hambaravi kihi kõikselt valimisse. Samuti parandab hinnangute täpsust kitsenduste kasutamine.

10. Kasutatud kirjandus

Knottnerus, P. (2003) Sampel Survey Theory. Some Pythagorean Perspectives. New York: Springer

Lepik, N, Traat, I. (2013) Valikuuringute teooria I. Aine konspekt.
http://www.e-ope.ee/_download/euni_repository/file/3991/ValiuurI.pdf

Lütkepohl, H. (1996) Handbook of Matrices. New York: Wiley

Sõstra, K. (2007) Restriction estimator for domains. Tartu, TÜ Kirjastus

Traat, I., Inno, J. (1997) Tõenäosuslik valikuuring, Tartu, TÜ Kirjastus

11. Lisad

Lisa 1.

Tabel 20. Ankeet

Asutuse KOOD		kood	1		
Isiku jrk. nr		arv	2		
Sugu		Kood	3		
Vanus		arv	4		
Ametiala		kood	5		
Juht		kood	6		
Arst-resident		kood	7		
Haridustase		kood	8		
Lepinguline koormus täistööaja suhtes		kood	9		
Tegelikult töötatud tunnid märtsis		tundi	10		
	sh ületunnid	tundi	11		
	sh valvetunnid	tundi	12		
Tööaja eest arvestatud töötasu märtsis		eurot	13		
	sh põhipalk	eurot	14		
	sh lisatasu öhtuse-, öötöö ja puhkep. riiklikel pühadel töötamise eest	eurot	15		
	sh tasu ületundide eest	eurot	16		
	sh tasu valvetundide eest	eurot	17		
	sh regulaarsed lisatasud	eurot	18		
Mitteregulaarne lisatasu		eurot	19		
Tasu mittetöötatud aja eest		eurot	20		

Lisa 2

Tabel 21. Töötajate arvud ja nende osakaalud TTO liigi järgi.

TTO liik	Arstid	Arstide osakaal	Õed	Õdede osakaal	Hooldajad	Hooldajate osakaal	Töötajaid kokku
Haigla	2 851	22%	6 072	47%	3 972	31%	12 895
Perearstiabi	798	43%	1 037	57%	0	0%	1 835
Eriarstiabi	511	55%	402	43%	15	2%	928
Hambaravi	2	10%	7	33%	12	57%	21
Kiirabi	91	22%	321	78%	0	0%	412
Taastusravi	49	26%	82	43%	58	31%	189
Muu	64	17%	301	79%	16	4%	381
Kokku	4 366	26%	8 222	49%	4 073	24%	16 661

Tabel 22. Töötajate arvud ja nende osakaalud haigla liigi järgi.

Haigla liik	Arstid	Arstide osakaal	Õed	Õdede osakaal	Hooldajad	Hooldajate osakaal	Töötajaid kokku
Piirkondlik	1 188	23%	2 440	47%	1 546	30%	5 174
Keskhaigla	862	23%	1 904	50%	1 034	27%	3 800
Üldhaigla	538	21%	1 152	46%	834	33%	2 524
Taastusravi haigla	13	13%	38	38%	49	49%	100
Hooldus-haigla	61	10%	235	40%	294	50%	590
Erihaigla	112	31%	174	48%	73	20%	359
Kohalik haigla	77	22%	129	37%	142	41%	348

Lisa 3

1.kood: LJV TTA hinnangu leidmine

```
/*LJV TTA valimi moodustamine, valimi maht= 30% ÜK mahust ehk
n=0,3*16661=4999*/
proc surveyselect data=b.uus
  reps=1000
  n=4999
  method=SRS
  out=b.LJVvalim;
run;
/*kaalu lisamine*/
data b.LJVvalim;
set b.LJVvalim;
kaal=3.332867;
run;
/*OSAKOGUMID*/
/*Ametigrupi järgi*/
/*arstide tunnipalga hindamine*/
data b.arstid;
set b.LJVvalim;
where ametigrupp = 1;
by replicate;
run;
proc surveymeans
  data=b.arstid
  total=16661 mean;
  weight kaal;
  var tunnipalk;
  by replicate;
  ods output Statistics=b.hinnang_arstid;
run;
proc means data=b.hinnang_arstid mean cv var;
var mean;
run;
/*õdede tunnipalga hindamine*/
data b.oed;
set b.LJVvalim;
where ametigrupp = 3;
by replicate;
run;
proc surveymeans
  data=b.oed
  total=16661 mean;
  weight kaal;
  var tunnipalk;
  by replicate;
  ods output Statistics=b.hinnang_oed;
run;
proc means data=b.hinnang_oed mean cv var;
var mean;
run;
/*Hooldajate tunnipalga hindamine*/
data b.hooldajad;
set b.LJVvalim;
where ametigrupp = 5;
by replicate;
run;
proc surveymeans
```

```

data=b.hooldajad
total=16661 mean;
weight kaal;
var tunnipalk;
by replicate;
ods output Statistics=b.hinnang_hooldajad;
run;
proc means data=b.hinnang_hooldajad mean cv var;
var mean;
run;
/*osakogumiteks on Tervishoiuteenuse osutajad + haiglad eraldi*/
/*eriarstiabi*/
data b.eriarstiabi;
set b.LJVvalim;
where teenus_uus = 25;
by replicate;
run;
proc surveymeans
data=b.eriarstiabi
total=16661 mean;
weight kaal;
var tunnipalk;
by replicate;
ods output Statistics=b.hinnang_eriarstiabi;
run;
proc means data=b.hinnang_eriarstiabi mean cv var;
var mean;
run;
/*perearstiabi*/
data b.perearstiabi;
set b.LJVvalim;
where teenus_uus = 20;
by replicate;
run;
proc surveymeans
data=b.perearstiabi
total=16661 mean;
weight kaal;
var tunnipalk;
by replicate;
ods output Statistics=b.hinnang_perearstiabi;
run;
proc means data=b.hinnang_perearstiabi mean cv var;
var mean;
run;
/*hambaravi*/
data b.hambaravi;
set b.LJVvalim;
where teenus_uus = 30;
by replicate;
run;
proc surveymeans
data=b.hambaravi
total=16661 mean;
weight kaal;
var tunnipalk;
by replicate;
ods output Statistics=b.hinnang_hambaravi;
run;
proc means data=b.hinnang_hambaravi mean cv var;
var mean;

```

```

run;
/*kiirabi*/
data b.kiirabi;
set b.LJVvalim;
where teenus_uus = 40;
by replicate;
run;
proc surveymeans
data=b.kiirabi
total=16661 mean;
weight kaal;
var tunnipalk;
by replicate;
ods output Statistics=b.hinnang_kiirabi;
run;
proc means data=b.hinnang_kiirabi mean cv var;
var mean;
run;
/*taastusravi*/
data b.taastusravi;
set b.LJVvalim;
where teenus_uus = 50;
by replicate;
run;
proc surveymeans
data=b.taastusravi
total=16661 mean;
weight kaal;
var tunnipalk;
by replicate;
ods output Statistics=b.hinnang_taastusravi;
run;
proc means data=b.hinnang_taastusravi mean cv var;
var mean;
run;
/*Muu*/
data b.ttomuu;
set b.LJVvalim;
where teenus_uus = 90;
by replicate;
run;
proc surveymeans
data=b.ttomuu
total=16661 mean;
weight kaal;
var tunnipalk;
by replicate;
ods output Statistics=b.hinnang_ttomuu;
run;
proc means data=b.hinnang_ttomuu mean cv var;
var mean;
run;
/*Haiglad*/
data b.ttohaiglad;
set b.LJVvalim;
where teenus_uus = 10;
by replicate;
run;
proc surveymeans
data=b.ttohaiglad
total=16661 mean;

```

```

weight kaal;
var tunnipalk;
by replicate;
ods output Statistics=b.hinnang_haiglad;
run;
proc means data=b.hinnang_haiglad mean cv var;
var mean;
run;
/*Haiglad eraldi liigi järgi*/
/*piirkondlik*/
data b.piirkondlik;
set b.LJVvalim;
where haigla_liik = 11;
by replicate;
run;
proc surveymeans
data=b.piirkondlik
total=16661 mean;
weight kaal;
var tunnipalk;
by replicate;
ods output Statistics=b.hinnang_piirkondlik;
run;
proc means data=b.hinnang_piirkondlik mean cv var;
var mean;
run;
/*keskhaigla*/
data b.keskhaigla;
set b.LJVvalim;
where haigla_liik = 12;
by replicate;
run;
proc surveymeans
data=b.keskhaigla
total=16661 mean;
weight kaal;
var tunnipalk;
by replicate;
ods output Statistics=b.hinnang_keskhaigla;
run;
proc means data=b.hinnang_keskhaigla mean cv var;
var mean;
run;
/*üldhaigla*/
data b.yldhaigla;
set b.LJVvalim;
where haigla_liik = 13;
by replicate;
run;
proc surveymeans
data=b.yldhaigla
total=16661 mean;
weight kaal;
var tunnipalk;
by replicate;
ods output Statistics=b.hinnang_yldhaigla;
run;
proc means data=b.hinnang_yldhaigla mean cv var;
var mean;
run;
/*taastusravihaigla*/

```

```

data b.taastusravi;
set b.LJVvalim;
where haigla_liik = 14;
by replicate;
run;
proc surveymeans
data=b.taastusravi
total=16661 mean;
weight kaal;
var tunnipalk;
by replicate;
ods output Statistics=b.hinnang_taastusravi;
run;
proc means data=b.hinnang_taastusravi mean cv var;
var mean;
run;
/*hooldushaigla*/
data b.hooldushaigla;
set b.LJVvalim;
where haigla_liik = 15;
by replicate;
run;
proc surveymeans
data=b.hooldushaigla
total=16661 mean;
weight kaal;
var tunnipalk;
by replicate;
ods output Statistics=b.hinnang_hooldushaigla;
run;
proc means data=b.hinnang_hooldushaigla mean cv var;
var mean;
run;
/*erihaigla*/
data b.erihaigla;
set b.LJVvalim;
where haigla_liik = 16;
by replicate;
run;
proc surveymeans
data=b.erihaigla
total=16661 mean;
weight kaal;
var tunnipalk;
by replicate;
ods output Statistics=b.hinnang_erihaigla;
run;
proc means data=b.hinnang_erihaigla mean cv var;
var mean;
run;
/*kohalik haigla*/
data b.kohalik;
set b.LJVvalim;
where haigla_liik = 17;
by replicate;
run;
proc surveymeans
data=b.kohalik
total=16661 mean;
weight kaal;
var tunnipalk;

```

```
by replicate;  
ods output Statistics=b.hinnang_kohalik;  
run;  
proc means data=b.hinnang_kohalik mean cv var;  
var mean;  
run;
```

Lisa 4

2.kood: PPS TGA hinnangu leidmine

```
/*PPS valimi moodustamine, valimi suurus 30% ÜK'st ehk 4999*/
proc surveyselect
data=b.uus
n=4999 /*valimi maht*/
method=pps_wr /*kasutatakse PPS TGA meetodit*/
reps=1000 /*võetakse valimit 1000 korda*/
out=b.ppsvalim; /*loodav andmestik*/
size haigla_suurus; /*taustunnuseks on haigla suurus*/
run;
/*Hindamine osakogumites*/
/*osakogumid AMETIGRUPPI järgi*/
/*arstid*/
data b.ppsarstid;
set b.ppsvalim;
where ametigrupp = 1; /*võtame andmestikust välja ainult arstid*/
by replicate; /*grupeeritakse valimite kaupa*/
run;
proc sql;
create table b.ppshin_arstid as
select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits) as
hinnangtunnipalk /*hinnangu leidmine*/
from b.ppsarstid
group by Replicate;
run;
quit;
proc means data=b.ppshin_arstid mean var cv;
var hinnangtunnipalk;
run;
/*õed*/
data b.ppsoed;
set b.ppsvalim;
where ametigrupp = 3;
by replicate;
run;
proc sql;
create table b.ppshin_oed as
select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits) as
hinnangtunnipalk
from b.ppsoed
group by Replicate;
run;
quit;
proc means data=b.ppshin_oed mean cv var;
var hinnangtunnipalk;
run;
/*hooldajad*/
data b.ppshooldajad;
set b.ppsvalim;
where ametigrupp = 5;
by replicate;
run;
proc sql;
create table b.ppshin_hooldajad as
```



```

select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits) as
hinnangtunnipalk
from b.ppshooldajad
group by Replicate;
run;
quit;
proc means data=b.ppshin_hooldajad mean cv var;
var hinnangtunnipalk;
run;
/*osakogumiteks on Tervishoiuteenuse osutajad + haiglad eraldi*/
/*haigla*/
data b.pps_haiglad;
set b.ppsvalim;
where teenus_uus = 10;
by replicate;
run;
proc sql;
create table b.ppshin_haiglad as
select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits) as
hinnangtunnipalk
from b.pps_haiglad
group by Replicate;
run;
quit;
proc means data=b.ppshin_haiglad mean cv var;
var hinnangtunnipalk;
run;
/*eriarstiabi*/
data b.pps_eriarstiabi;
set b.ppsvalim;
where teenus_uus = 25;
by replicate;
run;
proc sql;
create table b.ppshin_eriarstiabi as
select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits) as
hinnangtunnipalk
from b.pps_eriarstiabi
group by Replicate;
run;
quit;
proc means data=b.ppshin_eriarstiabi mean cv var;
var hinnangtunnipalk;
run;
/*perearstiabi*/
data b.ppsperearstiabi;
set b.ppsvalim;
where teenus_uus = 20;
by replicate;
run;
proc sql;
create table b.ppshin_perearstiabi as
select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits) as
hinnangtunnipalk
from b.ppsperearstiabi
group by Replicate;
run;

```

```

quit;
proc means data=b.ppshin_perearstiabi mean cv var;
var hinnangtunnipalk;
run;
/*hambaravi*/
data b.ppshambaravi;
set b.ppsvalim;
where teenus_uus = 30;
by replicate;
run;
proc sql;
create table b.ppshin_hambaravi as
select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits) as
hinnangtunnipalk
from b.ppshambaravi
group by Replicate;
run;
quit;
proc means data=b.ppshin_hambaravi mean cv var;
var hinnangtunnipalk;
run;
/*kiirabi*/
data b.ppskiirabi;
set b.ppsvalim;
where teenus_uus = 40;
by replicate;
run;
proc sql;
create table b.ppshin_kiirabi as
select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits) as
hinnangtunnipalk
from b.ppskiirabi
group by Replicate;
run;
quit;
proc means data=b.ppshin_kiirabi mean cv var;
var hinnangtunnipalk;
run;
/*taastusravi*/
data b.ppstaastusravi;
set b.ppsvalim;
where teenus_uus = 50;
by replicate;
run;
proc sql;
create table b.ppshin_taastusravi as
select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits) as
hinnangtunnipalk
from b.ppstaastusravi
group by Replicate;
run;
quit;
proc means data=b.ppshin_taastusravi mean cv var;
var hinnangtunnipalk;
run;
/*Muu*/
data b.ppsttomuu;
set b.ppsvalim;

```

```

where teenus_uus = 90;
by replicate;
run;
proc sql;
create table b.ppshin_ttomuu as
select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits) as
hinnangtunnipalk
from b.ppsttomuu
group by Replicate;
run;
quit;
proc means data=b.ppshin_ttomuu mean cv var;
var hinnangtunnipalk;
run;
/*Haiglad eraldi liigi järgi*/
/*piirkondlik*/
data b.ppspiirkondlik;
set b.ppsvalim;
where haigla_liik = 11;
by replicate;
run;
proc sql;
create table b.ppshin_piirkondlik as
select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits) as
hinnangtunnipalk
from b.ppspiirkondlik
group by Replicate;
run;
quit;
proc means data=b.ppshin_piirkondlik mean cv var;
var hinnangtunnipalk;
run;
/*keskhaigla*/
data b.ppskeskhaigla;
set b.ppsvalim;
where haigla_liik = 12;
by replicate;
run;
proc sql;
create table b.ppshin_keskhaigla as
select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits) as
hinnangtunnipalk
from b.ppskeskhaigla
group by Replicate;
run;
quit;
proc means data=b.ppshin_keskhaigla mean cv var;
var hinnangtunnipalk;
run;
/*üldhaigla*/
data b.ppsyldhaigla;
set b.ppsvalim;
where haigla_liik = 13;
by replicate;
run;
proc sql;
create table b.ppshin_yldhaigla as

```

```

select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits) as
hinnangtunnipalk
from b.ppsyldhaigla
group by Replicate;
run;
quit;
proc means data=b.ppshin_yldhaigla mean cv var;
var hinnangtunnipalk;
run;
/*taastusravihaigla*/
data b.ppstaastusravi;
set b.ppsvalim;
where haigla_liik = 14;
by replicate;
run;
proc sql;
create table b.ppshin_taastusr as
select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits) as
hinnangtunnipalk
from b.ppstaastusravi
group by Replicate;
run;
quit;
proc means data=b.ppshin_taastusr mean cv var;
var hinnangtunnipalk;
run;
/*hooldushaigla*/
data b.ppshooldush;
set b.ppsvalim;
where haigla_liik = 15;
by replicate;
run;
proc sql;
create table b.ppshin_hooldush as
select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits) as
hinnangtunnipalk
from b.ppshooldush
group by Replicate;
run;
quit;
proc means data=b.ppshin_hooldush mean cv var;
var hinnangtunnipalk;
run;
/*erihaigla*/
data b.ppserih;
set b.ppsvalim;
where haigla_liik = 16;
by replicate;
run;
proc sql;
create table b.ppshin_erih as
select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits) as
hinnangtunnipalk
from b.ppserih
group by Replicate;
run;
quit;

```

```

proc means data=b.ppshin_erih mean cv var;
var hinnangtunnipalk;
run;
/*kohalik haigla*/
data b.ppskohalik;
set b.ppsvalim;
where haigla_liik = 17;
by replicate;
run;
proc sql;
create table b.ppshin_kohalik as
select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits) as
hinnangtunnipalk
from b.ppskohalik
group by Replicate;
run;
quit;
proc means data=b.ppshin_kohalik mean cv var;
var hinnangtunnipalk;
run;

```

Lisa 5

3.kood: Lihtne juhuslik kihtvalik

```
/* Kihtidest võetakse valimid LJV TTA disaini kasutades.*/  
proc surveyselect data=b.uus rep=1000  
method=SRS  
sampsize=(3869 550 278 7 124 57 114)  
out=b.KLJV;  
strata teenus_uus;  
run;  
/* Kiht = 10 ehk haiglad*/  
data haiglate;  
set b.KLJV;  
if teenus_uus = 10 then output;  
run;  
proc sql;  
create table haiglate_hinnang as  
select sum(tunnipalk*SamplingWeight)/12895 as hinnangtunnipalk  
from haiglate  
group by Replicate;  
run;  
quit;  
proc means data=haiglate_hinnang mean cv var;  
var hinnangtunnipalk;  
run;  
/* Kiht = 20 ehk perearstiabi*/  
data perearstiabi;  
set b.KLJV;  
if teenus_uus = 20 then output;  
run;  
proc sql;  
create table perearstiabi_hinnang as  
select sum(tunnipalk*SamplingWeight)/1835 as hinnangtunnipalk  
from perearstiabi  
group by Replicate;  
run;  
quit;  
proc means data=perearstiabi_hinnang mean cv var;  
var hinnangtunnipalk;  
run;  
/* Kiht = 25 ehk eriarstiabi*/  
data eriarstiabi;  
set b.KLJV;  
if teenus_uus = 25 then output;  
run;  
proc sql;  
create table eriarstiabi_hinnang as  
select sum(tunnipalk*SamplingWeight)/928 as hinnangtunnipalk  
from eriarstiabi  
group by Replicate;  
run;  
quit;  
proc means data=eriarstiabi_hinnang mean cv var;  
var hinnangtunnipalk;  
run;  
/* Kiht = 30 ehk hambaravi*/  
data hambaravi;  
set b.KLJV;  
if teenus_uus = 30 then output;
```

```

run;
proc sql;
create table hambaravi_hinnang as
select sum(tunnipalk*SamplingWeight)/21 as hinnangtunnipalk
from hambaravi
group by Replicate;
run;
quit;
proc means data=hambaravi_hinnang mean cv var;
var hinnangtunnipalk;
run;
/* Kiht = 40 ehk kiirabi*/
data kiirabi;
set b.KLJV;
if teenus_uus = 40 then output;
run;
proc sql;
create table kiirabi_hinnang as
select sum(tunnipalk*SamplingWeight)/412 as hinnangtunnipalk
from kiirabi
group by Replicate;
run;
quit;
proc means data=kiirabi_hinnang mean cv var;
var hinnangtunnipalk;
run;
/* Kiht = 50 ehk taastusravi*/
data taastusravi;
set b.KLJV;
if teenus_uus = 50 then output;
run;
proc sql;
create table taastusravi_hinnang as
select sum(tunnipalk*SamplingWeight)/189 as hinnangtunnipalk
from taastusravi
group by Replicate;
run;
quit;
proc means data=taastusravi_hinnang mean cv var;
var hinnangtunnipalk;
run;
/* Kiht = 90 ehk muu*/
data muu;
set b.KLJV;
if teenus_uus = 90 then output;
run;
proc sql;
create table muu_hinnang as
select sum(tunnipalk*SamplingWeight)/381 as hinnangtunnipalk
from muu
group by Replicate;
run;
quit;
proc means data=muu_hinnang mean cv var;
var hinnangtunnipalk;
run;
/*Haiglate kihis moodustame eraldi haigla liigi järgi osakogumid ja leiame
nendes osakogumites hinnangud*/
/*piirkondlik*/
data piirkondlik;
set haiglate;

```

```

where haigla_liik = 11;
by replicate;
run;
proc surveymeans
data=piirkondlik
total=16661 mean;
weight SamplingWeight;
var tunnipalk;
by replicate;
ods output Statistics=hinnang_piirkondlik;
run;
proc means data=hinnang_piirkondlik mean cv var;
var mean;
run;
/*keskhaigla*/
data keskhaigla;
set b.KLJV;
where haigla_liik = 12;
by replicate;
run;
proc surveymeans
data=keskhaigla
total=16661 mean;
weight SamplingWeight;
var tunnipalk;
by replicate;
ods output Statistics=hinnang_keskhaigla;
run;
proc means data=hinnang_keskhaigla mean cv var;
var mean;
run;
/*üldhaigla*/
data yldhaigla;
set b.KLJV;
where haigla_liik = 13;
by replicate;
run;
proc surveymeans
data=yldhaigla
total=16661 mean;
weight SamplingWeight;
var tunnipalk;
by replicate;
ods output Statistics=hinnang_yldhaigla;
run;
proc means data=hinnang_yldhaigla mean cv var;
var mean;
run;
/*taastusravihaigla*/
data taastusravihaigla;
set b.KLJV;
where haigla_liik = 14;
by replicate;
run;
proc surveymeans
data=taastusravihaigla
total=16661 mean;
weight SamplingWeight;
var tunnipalk;
by replicate;
ods output Statistics=hinnang_taastusravihaigla;

```



```

run;
proc means data=hinnang_taastusravihaigla mean cv var;
var mean;
run;
/*hooldushaigla*/
data hooldushaigla;
set b.KLJV;
where haigla_liik = 15;
by replicate;
run;
proc surveymeans
data=hooldushaigla
total=16661 mean;
weight SamplingWeight;
var tunnipalk;
by replicate;
ods output Statistics=hinnang_hooldushaigla;
run;
proc means data=hinnang_hooldushaigla mean cv var;
var mean;
run;
/*erihaigla*/
data erihaigla;
set b.KLJV;
where haigla_liik = 16;
by replicate;
run;
proc surveymeans
data=erihaigla
total=16661 mean;
weight SamplingWeight;
var tunnipalk;
by replicate;
ods output Statistics=hinnang_erihaigla;
run;
proc means data=hinnang_erihaigla mean cv var;
var mean;
run;
/*kohalik haigla*/
data khaigla;
set b.KLJV;
where haigla_liik = 17;
by replicate;
run;
proc surveymeans
data=khaigla
total=16661 mean;
weight SamplingWeight;
var tunnipalk;
by replicate;
ods output Statistics=hinnang_khaigla;
run;
proc means data=hinnang_khaigla mean cv var;
var mean;
run;
/*Leiame saadud valimis erinevates ametigruppides hinnangud*/
/*Osakogumiks - arstid*/
data arstid;
set b.KLJV;
where ametigrupp = 1;
by replicate;

```

```

run;
proc surveymeans
data=arstid
total=16661 mean;
weight SamplingWeight;
var tunnipalk;
by replicate;
ods output Statistics=hinnang_arstid;
run;
proc means data=hinnang_arstid mean cv var;
var mean;
run;
/*Osakogumiks - õed*/
data oed;
set b.KLJV;
where ametigrupp = 3;
by replicate;
run;
proc surveymeans
data=oed
total=16661 mean;
weight SamplingWeight;
var tunnipalk;
by replicate;
ods output Statistics=hinnang_oed;
run;
proc means data=hinnang_oed mean cv var;
var mean;
run;
/*osakogumiks - hooldajad*/
data hooldajad;
set b.KLJV;
where ametigrupp = 5;
by replicate;
run;
proc surveymeans
data=hooldajad
total=16661 mean;
weight SamplingWeight;
var tunnipalk;
by replicate;
ods output Statistics=hinnang_hooldajad;
run;
proc means data=hinnang_hooldajad mean cv var;
var mean;
run;

```

Lisa 6

4.kood: PPS TGA kihtvalik

```
/*kihtideks tegemine */
/*sorteerime andmestiku tunnuse järgi, mille abil kihid moodustame*/
proc sort data=b.uus;
by teenus_uus;
run;
/*teatud haigla liigid ja hambaravi satuvad kõiksest valimisse, võtan
andmestikust välja*/
data b.uus_kind_haiglaliigid_out;
set b.uus;
if haigla_liik = 14 or haigla_liik = 15 or haigla_liik = 16 or haigla_liik
= 17 or teenus_uus = 30 then delete;
run;
proc sort data= b.uus_kind_haiglaliigid_out;
by teenus_uus haigla_suurus;
run;
/*Valimi jaoks jääb 21,5% ÜK mahust, siis valimi suuruseks on 3582.*/
proc surveyselect data=b.uus_kind_haiglaliigid_out rep=1000
method=pps_wr
samsize=(2701, 431, 218,97,45,90)
out=b.kihid_inim;
strata teenus_uus;
size haigla_suurus;
run;
/* Kiht = 20 ehk perearstiabi*/
data b.kihtperearstiabi;
set b.kihid_inim;
if teenus_uus = 20 then output;
run;
proc sql;
create table b.kihtperearstiabi_hinnang as
select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits) as
hinnangtunnipalk
from b.kihtperearstiabi
group by Replicate;
run;
quit;
proc means data=b.kihtperearstiabi_hinnang mean cv var;
var hinnangtunnipalk;
run;
/* Kiht = 25 ehk eriarstiabi*/
data b.kihteriarstiabi;
set b.kihid_inim;
if teenus_uus = 25 then output;
run;
proc sql;
create table b.kihteriarstiabi_hinnang as
select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits) as
hinnangtunnipalk
from b.kihteriarstiabi
group by Replicate;
run;
quit;
proc means data=b.kihteriarstiabi_hinnang mean cv var;
var hinnangtunnipalk;
run;
```

```

/*Hambaravil tuleb keskmine*/
data b.hambaravi;
set b.uus;
if teenus_uus = 30 then output;
run;
proc means data = b.hambaravi sum mean;
var tunnipalk;
run;
/* Kiht = 40 ehk kiirabi*/
data b.kihtkiirabi;
set b.kihid_inim;
if teenus_uus = 40 then output;
run;
proc sql;
create table b.kihtkiirabi_hinnang as
select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits) as
hinnangtunnipalk
from b.kihtkiirabi
group by Replicate;
run;
quit;
proc means data=b.kihtkiirabi_hinnang mean cv var;
var hinnangtunnipalk;
run;
/* Kiht = 50 ehk travi*/
data b.kihttravi;
set b.kihid_inim;
if teenus_uus = 50 then output;
run;
proc sql;
create table b.kihttravi_hinnang as
select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits) as
hinnangtunnipalk
from b.kihttravi
group by Replicate;
run;
quit;
proc means data=b.kihttravi_hinnang mean cv var;
var hinnangtunnipalk;
run;
/* Kiht = 90 ehk travi*/
data b.kihtmuu;
set b.kihid_inim;
if teenus_uus = 90 then output;
run;
proc sql;
create table b.kihtmuu_hinnang as
select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits) as
hinnangtunnipalk
from b.kihtmuu
group by Replicate;
run;
quit;
proc means data=b.kihtmuu_hinnang mean cv var;
var hinnangtunnipalk;
run;
/*osakogumid haigla kihis*/
/*haigla_liik = 11 ehk piirkondlik*/

```

```

data b.kihtpiirkondlik;
set b.kihid_inim;
if haigla_liik = 11 then output;
run;
proc sql;
create table b.kihtpiirkondlik_hinnang as
select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits) as
hinnangtunnipalk
from b.kihtpiirkondlik
group by Replicate;
run;
quit;
proc means data=b.kihtpiirkondlik_hinnang mean cv var;
var hinnangtunnipalk;
run;
/*haigla_liik = 12 ehk keskhaigla*/
data b.kihtkeskh;
set b.kihid_inim;
if haigla_liik = 12 then output;
run;
proc sql;
create table b.kihtkeskh_hinnang as
select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits) as
hinnangtunnipalk
from b.kihtkeskh
group by Replicate;
run;
quit;
proc means data=b.kihtkeskh_hinnang mean cv var;
var hinnangtunnipalk;
run;
/*haigla_liik = 13 ehk üldhaigla*/
data b.kihtyldh;
set b.kihid_inim;
if haigla_liik = 13 then output;
run;
proc sql;
create table b.kihtyldh_hinnang as
select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits) as
hinnangtunnipalk
from b.kihtyldh
group by Replicate;
run;
quit;
proc means data=b.kihtyldh_hinnang mean cv var;
var hinnangtunnipalk;
run;
/*Arstide, õdede ja hooldajate osakogumites hinnangud*/
/*leiam kogusumma hinnangud*/
/*ametigrupp = 1 ehk arstid*/
data b.kihtarstid;
set b.kihid_inim;
if ametigrupp = 1 then output;
run;
proc sql;
create table b.kihtarstid_hinnang as

```

```

select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits)*436
6 as hinnangtunnipalk
from b.kihtarstid
group by Replicate;
run;
quit;
proc means data=b.kihtarstid_hinnang mean cv va;
var hinnangtunnipalk;
run;
/*ametigrupp = 3 ehk oed*/
data b.kihtoed;
set b.kihid_inim;
if ametigrupp = 3 then output;
run;
proc sql;
create table b.kihtoed_hinnang as
select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits)*822
2 as hinnangtunnipalk
from b.kihtoed
group by Replicate;
run;
quit;
proc means data=b.kihtoed_hinnang mean cv var;
var hinnangtunnipalk;
run;
/*ametigrupp = 5 ehk hooldajad*/
data b.kihthooldajad;
set b.kihid_inim;
if ametigrupp = 5 then output;
run;
proc sql;
create table b.kihthooldajad_hinnang as
select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits)*407
3 as hinnangtunnipalk
from b.kihthooldajad
group by Replicate;
run;
quit;
proc means data=b.kihthooldajad_hinnang mean cv var;
var hinnangtunnipalk;
run;
/*Leiame põhitunnipalga kogusummad selles osas valimis, mis oli kõikne*/
data b.koikne;
set b.uus;
if haigla_liik = 14 or haigla_liik = 15 or haigla_liik = 16 or haigla_liik
= 17 or teenus_uus = 30 then output;
run;
proc sort data = b.koikne;
by ametigrupp;
run;
proc means data = b.koikne sum;
var tunnipalk;
by ametigrupp;
run;

```

Lisa 7

5.kood: Kitsendustega hinnang

```
/*tto_liikide kogusumma hinnang*/
/*Perëarstiabi kogusumma hinnang*/
proc sql;
create table b.kogusumma1 as
select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits)*183
5 as kogusumma_1, replicate
from b.kihtperëarstiabi
group by Replicate;
run;
quit;
proc means data=b.kogusumma1 mean var cv;
var kogusumma_1;
run;
/*Eriarstiabi kogusumma hinnang*/
proc sql;
create table b.kogusumma2 as
select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits)*928
as kogusumma_2, replicate
from b.kihteriarstiabi
group by Replicate;
run;
quit;
proc means data=b.kogusumma2 mean var cv;
var kogusumma_2;
run;
/*Kiirabi kogusumma hinnang*/
proc sql;
create table b.kogusumma3 as
select
replicate,sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*Numbe
rHits)*412 as kogusumma_3
from b.kihtkiirabi
group by Replicate;
run;
quit;
proc means data=b.kogusumma3 mean var cv;
var kogusumma_3;
run;
/*Taastusravi kogusumma hinnang*/
proc sql;
create table b.kogusumma4 as
select
replicate,sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*Numbe
rHits)*189 as kogusumma_4
from b.kihttravi
group by Replicate;
run;
quit;
proc means data=b.kogusumma4 mean var cv;
var kogusumma_4;
run;
/*Muu kogusumma hinnang*/
proc sql;
create table b.kogusumma5 as
```

```

select replicate,
(sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits))*3
81 as kogusumma_5
from b.kihtmuu
group by Replicate;
run;
quit;
proc means data= b.kogusumma5 mean var cv;
var kogusumma_5;
run;
data b.tto_kogusumma_ilmahaiglata;
merge b.kogusumma1 b.kogusumma2 b.kogusumma3 b.kogusumma4 b.kogusumma5;
by replicate;
run;
/*haigla kogusumma hinnang*/
proc sql;
create table kogusumma_haigla1 as
select replicate,
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits)*517
4 as kogusumma1 from b.kihtpiirkondlik
group by Replicate;
run;
quit;
proc means data = kogusumma_haigla1 mean var cv;
var kogusumma1;
run;
proc sql;
create table kogusumma_haigla2 as
select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits)*380
0 as kogusumma2, replicate from b.kihtkeskh
group by Replicate;
run;
quit;
proc means data = kogusumma_haigla2 mean var cv;
var kogusumma2;
run;
proc sql;
create table kogusumma_haigla3 as
select
sum(tunnipalk*SamplingWeight*NumberHits)/sum(Samplingweight*NumberHits)*252
4 as kogusumma3, replicate from b.kihtyldh
group by Replicate;
run;
quit;
proc means data = kogusumma_haigla3 mean var cv;
var kogusumma3;
run;
data haiglakogusumma_hinnangud;
merge kogusumma_haigla1 kogusumma_haigla2 kogusumma_haigla3;
by replicate;
run;
data b.haigla_kogusumma;
set haiglakogusumma_hinnangud;
kogusumma_hinnang = kogusumma1+kogusumma2+kogusumma3+6757.2238;
run;
proc means data = b.haigla_kogusumma mean var cv;
var kogusumma_hinnang;
run;

data b.haiglate_kogusummad;

```



```
merge b.tto_kogusumma_ilmahaiglata b.haigla_kogusumma;
by replicate;
run;

proc sql;
create table b.tto_kogusummahinnangud
as select kogusumma_1, kogusumma_2, kogusumma_3, kogusumma_4, kogusumma_5,
kogusumma_hinnang from b.haiglate_kogusummad;
run;
quit;
```

Kitsenduste rakendamine, kood R's.

#Loeme sisse kovaratsiooni maatriksi V

```
V = matrix(c(151310.456,-897.8294, 392.4296,673.5285,1411.6980,-9286.6520,
-897.8294,1428630.42,3176.1874,4231.6443,-17187.171,-21299.598,
392.4296,3176.1874,5087.6605,177.4877,66.7064,-607.7914,
673.5285,4231.6443,177.4877, 12610.7957,544.3714,5483.1823,
1411.6980,-17187.171,66.7064,544.3714,41645.1709, 2144.3053,
-9286.6520,-21299.598,-607.7914,5483.1823,2144.3053,1545722.68),nrow=6, ncol=6)
```

V

#1000 simuleeritud valimist saadud andmete sisselugemine

```
a=read.table("C:/Users/Kristi/Desktop/KOOL/Loputoo/b/tabel3.csv",sep = ",", header = T)
a=as.matrix(a)
a=t(a)
c=92210.9828 # c-c1
```

```
R=t(rep(1,6))
```

```
I=diag(6)
```

```
RVRt=R%*%V%*%t(R)
```

```
K=V%*%t(R)*(RVRt[1])^(-1)
```

```
hinnangud = matrix(rep(NA,6000), nrow=6)
```

```
for (i in 1:1000) {
```

```
deeta = a[,i]
```

```
hinnangud[,i] = deeta + K%*%(c-R%*%deeta)
```

```
}
```

```
hinnangud[,1:6]
```

```
hinnang = rowMeans(hinnangud)
```

```

hinnang
var(hinnangud[1,])
var(hinnangud[2,])
var(hinnangud[3,])
var(hinnangud[4,])
var(hinnangud[5,])
var(hinnangud[6,])
#Leian suhtelise vea protsentides
suht1=sqrt(var(hinnangud[1,]))/hinnang[1]*100
suht2=sqrt(var(hinnangud[2,]))/hinnang[2]*100
suht3=sqrt(var(hinnangud[3,]))/hinnang[3]*100
suht4=sqrt(var(hinnangud[4,]))/hinnang[4]*100
suht5=sqrt(var(hinnangud[5,]))/hinnang[5]*100
suht6=sqrt(var(hinnangud[6,]))/hinnang[6]*100
suht1
suht2
suht3
suht4
suht5
suht6

```

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina, Kristi Tüli,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) enda loodud teose „Võrdeliste töönaosustega valik tervishoiutöötajate tunnipalga uurimiseks“, mille juhendaja on Natalja Lepik,
 - 1.1.reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;
 - 1.2.üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace'i kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.
2. olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.
3. kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tartus, 05.05.2014